



TESIS - TE2599

**SISTEM INFORMASI GEOGRAFIS (SIG)
SEBARAN DAERAH ALIRAN SUNGAI (DAS)
MENGUNAKAN SEGMENTASI BERBASIS
ALGORITMA *K-MODES CLUSTERING* DAN
*DAVIES-BOULDIN INDEX***

**MUHAMMAD FARID FAHMI
2214206704**

**DOSEN PEMBIMBING
Dr. Ir. Yoyon Kusnendar Suprpto, M.Sc.
Dr. Ir. Wirawan, DEA.**

**PROGRAM MAGISTER
BIDANG KEAHLIAN TELEMATIKA - CIO
JURUSAN TEKNIK ELEKTRO
FAKULTAS TEKNOLOGI INDUSTRI
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2016**



TESIS - TE2599

**GEOGRAPHIC INFORMATION SYSTEM (GIS)
DISTRIBUTION OF WATERSHED USING
SEGMENTATION BASED OF *K-MODES*
CLUSTERING ALGORITHM AND *DAVIES-
BOULDIN INDEX***

MUHAMMAD FARID FAHMI
2214206704

Supervisor
Dr. Ir. Yoyon Kusnendar Suprpto, M.Sc.
Dr. Ir. Wirawan, DEA.

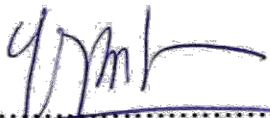
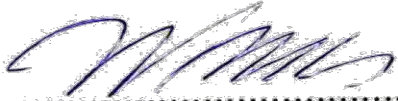
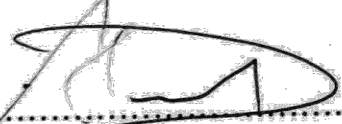
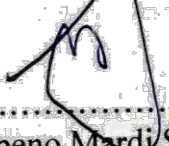
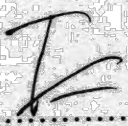
MASTER PROGRAME OF
TELEMATICS - CHIEF INFORMATION OFFICER
DEPARTEMEN OF ELECTRICAL ENGINEERING
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2016

LEMBAR PENGESAHAN

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Teknik (M.T)
di
Institut Teknologi Sepuluh Nopember
oleh ;
Muhammad Farid Fahmi
Nrp. 2214206704

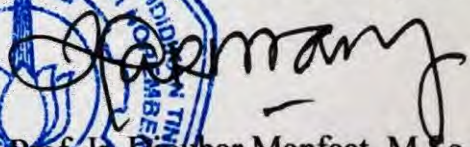
Tanggal Ujian : 23 Juni 2016
Periode Wisuda : Wisuda ke-114

Disetujui oleh :

1. 
.....
Dr. Ir. Yoyon K Suprpto, M.Sc.
NIP. 19540925 197803 1001
(Pembimbing I)
2. 
.....
Dr. Ir. Wirawan, DEA.
NIP. 19631109 198903 1011
(Pembimbing II)
3. 
.....
Mochamad Hariadi, ST., M.Sc., P.hD.
NIP. 19691209 199703 1002
(Penguji)
4. 
.....
Dr. Supeno Mardi Susiki, ST., MT.
NIP. 19700313 199512 1001
(Penguji)
5. 
.....
Dr. Istas Pratomo, ST., MT.
NIP. 19790325 200312 1001
(Penguji)



Direktur Program Pasca Sarjana,


Prof. Ir. Djouhar Manfaat, M.Sc, Ph.D
NIP. 19601202 198701 1001

SISTEM INFORMASI GEOGRAFIS (SIG) SEBARAN DAERAH ALIRAN SUNGAI (DAS) MENGGUNAKAN SEGMENTASI BERBASIS ALGORITMA K-MODES CLUSTERING DAN DAVIES-BOULDIN INDEX

Nama : Muhammad Farid Fahmi
NRP : 2214206704
Pembimbing I : Dr. Ir. Yoyon K Suprpto, M.Sc
Pembimbing II : Dr. Ir. Wirawan, DEA

ABSTRAK

Tingkat keberhasilan rehabilitasi DAS saat ini masih belum maksimal, salah satu penyebabnya adalah keterbatasan informasi tentang kondisi DAS. Dari permasalahan di atas, diperlukan suatu penelitian yang dapat memberikan acuan atau alternatif lain dalam menentukan DAS prioritas untuk direhabilitasi, salah satunya melalui data mining. Dalam penelitian ini DAS akan dikelompokkan menggunakan algoritma K-modes clustering berdasarkan parameter karakteristiknya. Hasil pengelompokan DAS dengan K-modes clustering kemudian dioptimalkan menggunakan Davies Bouildin Index (DBI) untuk mendapatkan jumlah cluster dengan tingkat kemiripan yang optimal dan dilakukan visualisasi dengan Sistem Informasi Geografis untuk memperoleh peta sebaran DAS. Dari uji coba pada DAS Tondano didapatkan bahwa cluster nomor empat (4) adalah jumlah cluster yang optimal dengan nilai DBI rata-rata 0,672778, atau 19,93%. Hasil clustering menunjukkan bahwa DAS dalam cluster 3 dengan 332 DAS yang sebagian besar tersebar di Minahasa Selatan (24,7%) adalah DAS kritis dibandingkan dengan kelompok lainnya. Hasil dari proses pengelompokan tidak jauh berbeda atau 90,64% sama jika dibandingkan dengan perhitungan DAS secara manual, yang dapat digunakan sebagai acuan atau alternatif lain dalam perencanaan rehabilitasi DAS.

Kata kunci: Data Mining, Clustering, K-Modes, Davies-Bouldin Index, SIG, DAS

GEOGRAPHIC INFORMATION SYSTEM (GIS) DISTRIBUTION OF WATERSHED USING SEGMENTATION BASED OF K-MODES CLUSTERING ALGORITHM AND DAVIES-BOULDIN INDEX

Name : Muhammad Farid Fahmi
NRP : 2214206704
Supervisor : Dr. Ir. Yoyon K Suprpto, M.Sc
Co Supervisor : Dr. Ir. Wirawan, DEA

ABSTRACT

The watershed rehabilitation success rate have not been up, is the result of policies in watershed rehabilitation strategies that are less precise. From the above problems, we need a study that can provide a reference or any other alternative in determining priority watersheds to be rehabilitated, one through data mining. This paper uses a case study of Watershed data which are grouped using K-modes clustering algorithm based on its characteristics parameters. Watershed grouped using K-modes clustering then optimized using Davies- Bouldin Index (DBI) to get the number of clusters with the optimal level of similarity and visualized using GIS to obtain distribution maps. From trial on the Watershed of Tondano It was known that the cluster number four (4) is the optimal cluster number with an average DBI value of 0.672778, or 19.93%. The clustering results show that the watershed in cluster 3 with 332 watershed which mostly scattered in the South Minahasa (24.7%) is a critical watershed compared to other clusters. the result of the clustering process is not much different or 90.64% similar when compared to the calculation of the watershed manually, that can be used as alternative to other reference in planning the rehabilitation of the watershed.

Keywords: Data Mining, Clustering, K-Modes, Davies-Bouldin Index, GIS, watershed

KATA PENGANTAR

Alhamdulillah, segala puji syukur penulis panjatkan kehadirat Allah SWT, atas segala karunia dan ridho-NYA, sehingga tesis dengan judul “*Sistem Informasi Geografis Sebaran Daerah Aliran Sungai (DAS) Menggunakan Segmentasi Berbasis Algoritma K-Modes Clustering dan Davies-Bouldin Index*” ini dapat diselesaikan.

Tesis ini disusun untuk memenuhi salah satu persyaratan memperoleh gelar Magister Teknik (M.T.) dalam bidang keahlian Telematika-CIO, Jurusan Teknik Elektro, Institut Teknologi Sepuluh Nopember Surabaya. Pada kesempatan ini penulis ingin menyampaikan rasa hormat dan ucapan terima kasih kepada:

1. Kementerian Komunikasi dan Informatika Melalui Beasiswa Dalam Negeri yang memberi kesempatan penulis untuk menempuh Pendidikan Magister Teknik Elektro Bidang Keahlian Telematika-CIO di Institut Teknologi Sepuluh Nopember Surabaya.
2. Orang tua tercinta Ayahanda Ainur Rouf, Ibunda Sholichah, dan semua saudara-saudaraku atas semua dukungan, bantuan dan doa-doa yang tidak pernah putus selama penulis belajar di S-2.
3. Istriku tercinta Astri Kurnia Nirwana Sari atas semua pengorbanan jiwa raga maupun kesabaran yang luar biasa selama mendampingi penulis menyelesaikan studi S-2, yang tidak akan pernah tergantikan oleh apapun. Mudah-mudahan Allah menjadikan pahala atas semua yang telah dilakukannya kepada penulis.
4. Anakku Rasya Muhammad Alvaro Fahmi atas semua rasa cinta dan sayang yang dihadirkan dalam kehidupan ini. Mudah-mudahan menjadi anak yang sholeh.
5. Bapak Dr. Ir. Yoyon Kusnendar Suprpto, M.Sc. atas bimbingan, arahan dan waktu yang telah diluangkan kepada penulis untuk berdiskusi selama menjadi dosen wali, dosen pembimbing dan perkuliahan.
6. Bapak Dr. Ir. Wirawan, DEA. atas bimbingan, arahan dan waktu yang telah diluangkan kepada penulis untuk berdiskusi selama menjadi dosen pembimbing dan perkuliahan.

7. Bapak Mochamad Hariadi, ST., M.Sc, P.hD, Bapak Dr. Supeno Mardi Susiki, ST.,MT serta bapak Dr. Istas Pratomo, ST.,MT atas masukan dan arahnya selama menjadi penguji dalam sidang tesis maupun sebagai dosen pengajar .
8. Bapak Dr. Ir. Djoko Purwanto, M.Eng. selaku Ketua Jurusan Teknik Elektro, yang memberikan ruang kepada penulis untuk mengembangkan diri dan berkarya.
9. Bapak Prof. Ir. Djauhar Manfaat, M.Sc, Ph.D. selaku Direktur Pasca Sarjana, yang memberikan kesempatan kepada penulis untuk mengembangkan diri dan berkarya di Institut Teknologi Sepuluh Nopember.
10. Semua Dosen-dosen Telematika-CIO ITS yang telah memberikan ilmu yang bermanfaat bagi penulis, mudah-mudahan dianggap sebagai amal jariyah oleh Allah SWT.
11. Rekan-rekan S2 Telematika-CIO dan S2 Telematika seangkatan maupun beda angkatan yang selalu memberikan motivasi maupun bantuan lainnya yang tidak bisa disebutkan satu persatu.
12. Kepala Balai Penelitian dan Pengembangan Lingkungan Hidup dan Kehutanan (BP2LHK) Manado Ir. Muh. Abidin, M.Si dan Rekan-rekan tempat penulis bekerja di BP2LHK manado atas semua dukungan dan doanya, sehingga penulis dapat menyelesaikan studi S-2.
13. Kepada semua pihak yang telah membantu, mendoakan, memberikan motivasi dan dorongan yang tidak dapat saya sebutkan satu persatu.

Surabaya, Juli 2016

Muhammad Farid Fahmi

DAFTAR ISI

HALAMAN SAMPUL	i
PERNYATAAN KEASLIAN TESIS	ii
LEMBAR PENGESAHAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xii
DAFTAR ISTILAH	xiii
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Manfaat Penelitian	3
1.5 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Daerah Aliran Sungai (DAS)	5
2.2 Rehabilitasi DAS	6
2.3 Parameter Karakteristik DAS	7
2.3.1 Penutupan Lahan	8
2.3.2 Tingkat Kemiringan Lereng	8
2.3.3 Tingkat Bahaya Erosi	9
2.3.4 Tingkat Produktivitas	9
2.3.5 Manajemen	9
2.4 Data Mining	10
2.5 Clustering	11
2.5.1 <i>Hierarical Clustering</i>	12

2.5.2 Non Hierarical Clustering	12
2.6 Analisis Cluster	13
2.7 Algoritma <i>K-Modes Clustering</i>	14
2.8 Validitas Cluster	16
2.8.1 Validitas Internal	16
2.8.2 Validitas Eksternal	17
2.8.3 <i>Davies Bouldin Index</i>	18
2.9 Sistem Informasi Geografis (SIG)	19
2.9.1 Data dan Operasi dalam SIG	20
2.9.2 Google Map API	21
2.10 Penelitian Terkait	22
BAB III METODOLOGI	25
3.1 Pra Prosessing Data	26
3.2 Clustering DAS dengan Algoritma <i>K-Modes Clustering</i>	26
3.3 Validitas Cluster dengan Metode <i>Davies Bouldin Index</i>	30
3.4 Analisis dan Optimasi Clustering	34
3.5 Visualisasi Cluster dengan Sistem Informasi Geografis (SIG)	35
BAB IV HASIL DAN PEMBAHASAN	37
4.1 Tahap Pra Prosesing Data	37
4.1.1 Transformasi Data	37
4.1.2 Pembersihan Data dan Pemilihan Data Uji	39
4.2 Hasil Clustering DAS dengan Algoritma <i>K-Modes Clustering</i>	40
4.3 Hasil Validitas Cluster dengan <i>Davies Bouldin Index</i>	43
4.4 Hasil Analisis dan Optimasi Clustering	46
4.5 Hasil Visualisasi Clustering DAS dengan Google Map API.....	53
BAB V KESIMPULAN	57
5.1 Kesimpulan	57
5.2 Saran	57
DAFTAR PUSTAKA	
LAMPIRAN	

DAFTAR GAMBAR

Gambar 2.1	Salah satu DAS di Kab. Bolaang Mongondow Utara	5
Gambar 2.2	Tahapan dalam Data Mining	10
Gambar 2.3	Ilustrasi data clustering	11
Gambar 2.4	Konsep Kohesi dan Separasi	17
Gambar 2.5	Contoh peta kota Manado berbasis Google Map API	21
Gambar 3.1	Diagram alur penelitian	25
Gambar 4.1	Konversi data dari format mdb ke MySQL menggunakan <i>MS Access to MySQL (v5.4)</i>	38
Gambar 4.2	Data berhasil dikonversi dari format .mdb ke MySQL menggunakan <i>MS Access to MySQL (v5.4)</i>	38
Gambar 4.3	Query database penentuan centroid awal secara random ...	40
Gambar 4.4	Centroid (modus) awal pada aplikasi	40
Gambar 4.5	Query untuk menghitung jarak data dengan centroid	41
Gambar 4.6	Jarak antara data dengan centorid dan cluster yang diikuti pada aplikasi	42
Gambar 4.7	Query untuk memperoleh centroid (modus) baru	42
Gambar 4.8	Centroid (modus) baru	43
Gambar 4.9	Jumlah data pada masing-masing cluster	43
Gambar 4.10	Query untuk menghitung nilai SSW dan SSB	44
Gambar 4.11	Nilai SSW dan SSB pada aplikasi	44
Gambar 4.12	Query untuk menghitung nilai R_{ij}	44
Gambar 4.13	Rekap nilai SSW, SSB dan R_{ij} pada aplikasi	45
Gambar 4.14	Query untuk menghitung nilai $R_{ij \max}$	45
Gambar 4.15	Nilai $R_{ij \max}$ pada aplikasi	45
Gambar 4.16	Proses clustering DAS secara keseluruhan	46
Gambar 4.17	Grafik nilai DBI rata-rata terhadap jumlah cluster	47
Gambar 4.18	Detail hasil clustering akhir DAS	49

Gambar 4.19 Koordinat Latitude dan Longitude Desa Malompar Utara, Kec. Belang, Kab. Minahasa Tenggara	53
Gambar 4.20 Data pada Tabel Plot_das	54
Gambar 4.21 Tampilan XML yang berisi data lokasi DAS.....	55
Gambar 4.22 Hasil marking semua lokasi DAS yang ditampilkan dalam Google Map	55
Gambar 4.23 Visualisasi clustering DAS Tondano berdasarkan parameter karakteristik	56

DAFTAR TABEL

Tabel 2.1	Skoring nilai kekritisian lahan	6
Tabel 2.2	Bobot dari parameter DAS untuk skoring kekritisian	7
Tabel 2.3.	Kategori dan Skoring Penutupan Lahan	8
Tabel 2.4.	Kategori dan Skoring Kemiringan Lereng	8
Tabel 2.5.	Kategori dan Skoring Tingkat Erosi	9
Tabel 2.6.	Kategori dan Skoring Tingkat Produktivitas	9
Tabel 2.7.	Kategori dan Skoring Manajemen	10
Tabel 2.8	Contoh data set kategorikal 3 dimensi	15
Tabel 3.1	Dataset 10 data DAS	26
Tabel 3.2	Centroid awal pada contoh data DAS	27
Tabel 3.3	Matrik jarak dan cluster yang diikuti pada iterasi ke-1 data DAS	28
Tabel 3.4	Data yang tergabung pada cluster 1 data DAS	28
Tabel 3.5	Data yang tergabung pada cluster 2 data DAS	29
Tabel 3.6	Data yang tergabung pada cluster 3 data DAS	29
Tabel 3.7	Centroid baru yang didapat	29
Tabel 3.8	Contoh Hasil Clustering data DAS	30
Tabel 3.9	Contoh centroid akhir Clustering data DAS	30
Tabel 3.10	Matrik nilai SSW contoh clustering data DAS	32
Tabel 3.11	Matrik nilai SSB contoh clustering data DAS	33
Tabel 3.12	Nilai R_{max} dan DBI contoh clustering data DAS	34
Tabel 4.1	Pembersihan dan Pemilihan data uji dengan <i>Query</i>	39
Tabel 4.2	Rekapitulasi nilai DBI	47
Tabel 4.3	Hasil clustering dengan 4 cluster	48
Tabel 4.4.	Karakteristik DAS pada setiap cluster/segmen	50
Tabel 4.5.	Sebaran DAS untuk Kab/Kota pada masing-masing cluster/segmen	51
Tabel 4.6	Perbandingan <i>jumlah DAS</i> hasil clustering dengan penghitungan manual tingkat kekritisian lahan	52

DAFTAR ISTILAH

DAS	: Daerah Aliran Sungai
RHL	: Rehabilitasi Hutan dan Lahan
BPDAS	: Balai Pengelolaan Daerah Aliran Sungai
SIG	: Sistem Informasi Geografis
TBE	: Tingkat Bahaya erosi
DBI	: Davies-Bouldin Index
$d(x, c)$: Jarak data x ke c
x_j	: Nilai fitur ke- j dari x
c_j	: Nilai fitur ke- j dari c
r	: Jumlah fitur dalam vector
SSW	: Sum of square within cluster
m_i	: Jumlah data dalam cluster i
SSB	: Sum of square between cluster
$d(c_i, c_j)$: Jarak centroid c_i dengan centroid c_j
R_{ij}	: Nilai Perbandingan cluster i dengan cluster j
K	: Jumlah cluster yang digunakan
API	: Application Programming Interface
SQL	: Structured Query Language

DAFTAR LAMPIRAN

- LAMPIRAN 1. Script koneksi database dengan PHP (koneksi.php).
- LAMPIRAN 2. Script XML sebagai perantara antara MySQL dengan Google Map (generatexml.php).
- LAMPIRAN 3. Script Menampilkan XML ke browser (script index.php).
- LAMPIRAN 4. Permohonan Ijin Penelitian

BAB I

PENDAHULUAN

1.1 Latar Belakang

Hutan Indonesia seluas 137,09 juta Ha atau $\pm 70\%$ dari luas daratan Indonesia adalah merupakan hutan tropis ketiga terbesar di dunia, mempunyai peran dan nilai yang sangat strategis. Sumberdaya hutan bukan hanya merupakan sumberdaya alam yang mempunyai nilai ekonomis, tetapi juga mempunyai nilai sosial, dan nilai ekologis. Mengingat luasnya, maka hutan Indonesia tidak hanya mempunyai nilai strategis bagi rakyat Indonesia tapi juga bagi lingkungan global, terutama dalam mengurangi dampak perubahan iklim dunia

Kerusakan sumber daya hutan berakibat pada menurunnya kemampuan hutan dalam mendukung fungsi ekonomi, sosial dan ekologis. Indikasi kerusakan sumber daya hutan ini dapat dilihat dari menurunnya kualitas Daerah Aliran Sungai (DAS) dan semakin intensnya terjadi bencana alam berupa banjir, kekeringan, dan tanah longsor. Selain itu, kerusakan sumber daya hutan menjadi sorotan dunia internasional sebagai salah satu penyebab perubahan iklim dunia.

Dalam rangka mengembalikan kondisi DAS, upaya Rehabilitasi hutan dan lahan (RHL) serta pengembangan fungsi Daerah Aliran Sungai terus ditingkatkan dan disempurnakan oleh pemerintah. Selama ini proses RHL dilakukan menggunakan acuan tingkat kekritisian lahan berdasarkan bobot parameter karakteristik DAS, dimana DAS dengan tingkat kekritisian lahan Sangat Kritis dan Kritis adalah DAS dengan prioritas utama yang akan direhabilitasi. Dari beberapa penelitian yang telah dilakukan, salah satunya di DAS Juwana ternyata didapatkan bahwa tingkat keberhasilan rehabilitasi DAS hanya 65,8% yang artinya masih belum maksimal (Surtiani, 2015).

Dari permasalahan diatas, diperlukan suatu penelitian yang mampu memberikan acuan atau alternatif lain dalam menentukan DAS prioritas yang akan direhabilitasi, salah satunya melalui teknik data mining yang bisa memberikan informasi tentang sebaran kelompok DAS berdasarkan parameter karakteristik DAS.

Salah satu teknik dalam data mining yang banyak diaplikasikan untuk membantu menyelesaikan berbagai permasalahan adalah *Clustering*. Dalam teknik *clustering* data akan dikelompokkan kedalam beberapa *cluster* berdasarkan kemiripan dari karakteristik data-data tersebut. pada penelitian ini algoritma clustering yang digunakan adalah *K-Modes Clustering*. *K-Modes Clustering* memodelkan dataset menjadi klaster-klaster dimana data pada satu klaster yang memiliki karakteristik sama akan terkelompok dan memiliki karakteristik yang berbeda dari klaster lain. Setelah data dikelompokkan dengan *K-Modes Clustering*, kemudian diolah dengan metode *Davies Bouldin Index* (DBI) untuk menghasilkan jumlah cluster dengan tingkat kemiripan yang terbaik.

Hasil cluster yang terbentuk selanjutnya akan dilakukan visualisasi berbasis Sistem Informasi Geografis (SIG) yang berfungsi untuk mempermudah dalam menganalisis sebaran kelompok-kelompok DAS berdasarkan parameter karakteristiknya. Sebaran DAS dapat dipetakan berdasarkan letak geografis dalam bentuk peta yang dipadukan dengan basisdata untuk menunjang informasi dari masing-masing DAS. Pengetahuan yang diperoleh dari visualisasi clustering diharapkan dapat dimanfaatkan sebagai bahan pertimbangan untuk menentukan prioritas DAS yang akan direhabilitasi dan strategi rehabilitasi yang akan dilakukan.

1.2 Rumusan masalah

Permasalahan yang diangkat dalam tesis ini adalah, penentuan prioritas DAS yang akan direhabilitasi selama ini masih menggunakan penghitungan manual tingkat kekritisian lahan sehingga diperlukan alternatif lain dalam menentukan DAS yang menjadi prioritas untuk direhabilitasi.

1.3 Tujuan penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah mengelompokkan DAS berdasarkan parameter karakteristik DAS menggunakan Algoritma *K-modes clustering* dan *Davies Bouldin Index* (DBI) serta melakukan visualisasi sebaran kelompok DAS berdasarkan karakteristiknya dengan Sistem Informasi Geografis (SIG).

1.4 Manfaat Penelitian

Sedangkan Manfaat dari penelitian ini antara lain memberikan informasi kepada organisasi dalam hal ini Kementerian Lingkungan Hidup dan Kehutanan tentang sebaran kelompok DAS berdasarkan karakteristiknya sehingga bisa membantu dalam menentukan prioritas DAS yang akan direhabilitasi dan strategi rehabilitasi yang akan dilakukan.

1.5 Sistematika Penulisan

Penulisan laporan hasil penelitian ini dibagi menjadi beberapa bab dan sub bab, Berikut akan dijelaskan mengenai sistematika penulisan laporan penelitian yang telah dilakukan :

Bab I Pendahuluan

Bab ini terdiri dari latar belakang dilakukannya penelitian ini kemudian dilanjutkan dengan perumusan masalah, tujuan dan manfaat penelitian dan sistematika penulisan laporan.

Bab II Tinjauan Pustaka

Pada bab dua akan dijabarkan beberapa kajian tentang penelitian terdahulu yang berhubungan dengan metode yang digunakan, definisi DAS dan parameter karakteristik DAS. Pada bab ini juga dijelaskan tentang data mining dengan k-modes clustering, validitas cluster dengan davies-bouldin index serta konsep Sistem Informasi Geografis.

Bab III Metodologi

Bab ini memberikan uraian tentang tahapan penelitian yang dilakukan, bagaimana melakukan clustering dan melakukan validasi hasil cluster serta bagaimana melakukan visualisasi hasil cluster dengan Sistem Informasi Geografis.

Bab IV Hasil dan Pembahasan

Pada bab ini dijabarkan implementasi clustering dan validasi cluster serta proses optimasi jumlah cluster. Hasil dari percobaan dan visualisasi dalam bentuk Sistem Informasi Geografis yang dilakukan dalam penelitian juga dijabarkan dalam bab ini.

Bab V Kesimpulan

Bab ini berisi hasil akhir yang menjawab tujuan penelitian ini. Bab ini juga menjelaskan tentang hasil klastering yang diperoleh, bagaimana anggota klaster dan intepretasi terhadap visualisasi hasil cluster yang dihasilkan.

BAB II

TINJAUAN PUSTAKA

2.1 Daerah Aliran Sungai (DAS)

Daerah Aliran Sungai (DAS) adalah suatu wilayah daratan yang merupakan satu kesatuan dengan sungai dan anak sungainya yang berfungsi menampung, menyimpan, dan mengalirkan air yang berasal dari curah hujan ke danau atau ke laut secara alami. yang batas di daratan merupakan pemisah topografis dan batas di laut sampai dengan daerah perairan yang masih terpengaruh aktivitas daratan (Kemenhut, 2009).

DAS yang tersebar diseluruh wilayah Indonesia, merupakan satu kesatuan ekosistem alami yang utuh dari ekosistem pegunungan di hulu hingga ekosistem pantai di hilir. Dalam prakteknya perencanaan dan pengelolaan DAS adalah upaya dalam mengelola hubungan timbal balik antara sumber daya alam dengan sumber daya manusia di dalam DAS dan segala aktivitasnya untuk mewujudkan kemanfaatan sumber daya alam bagi kepentingan pembangunan dan kelestarian ekosistem DAS serta kesejahteraan masyarakat . Gambaran salah satu DAS di tunjukan pada Gambar 2.1.



Sumber : BP2LHK Manado

Gambar 2.1. Salah satu DAS di Kab. Bolaang Mongondow Utara

2.2 Rehabilitasi DAS

Rehabilitasi Daerah Aliran Sungai atau DAS adalah upaya untuk memulihkan, mempertahankan dan meningkatkan fungsi DAS sehingga daya dukung DAS terhadap fungsi kawasan hutan tetap terjaga. Tujuan kegiatan rehabilitasi dan penanganan DAS adalah tercapainya optimalisasi rehabilitasi lahan kritis, peningkatan produktivitas hutan dan pemanfaatan hutan/lahan (Kemenhut, 2009).

Selama ini proses RHL pada DAS dilakukan menggunakan acuan tingkat kekritisian lahan berdasarkan bobot parameter karakteristik DAS, dimana DAS dengan tingkat kekritisian lahan Sangat Kritis dan Kritis adalah DAS dengan prioritas utama yang akan direhabilitasi. Metode penilaian lahan kritis mengacu pada definisi lahan kritis yaitu sebagai lahan yang telah mengalami kerusakan.

Untuk masing-masing fungsi lahan, ditentukan parameter DAS yang terbagi lagi kedalam beberapa kelas. Untuk penilaiannya, pada masing-masing kelas diberi skoring serta bobot dari parameter DAS. Jumlah total skor dikalikan bobot dari masing-masing parameter DAS merupakan kategori kekritisian lahan masing-masing kawasan. Kategori skoring nilai kekritisian lahan di Kawasan Hutan ditunjukkan pada Tabel 2.1.

Tabel 2.1 Skoring nilai kekritisian lahan

Kawasan Hutan		Lindung	Budidaya	Luar kawasan
No.	Tingkat Kekritisian Lahan	Skor	Skor	Skor
1.	Sangat kritis	120 – 180	115 – 200	110 – 200
2.	Kritis	181 – 270	201 – 275	201 – 275
3.	Agak kritis	271 – 360	276 – 350	276 – 350
4.	Potensial kritis	361 – 450	351 – 425	351 – 425
5.	Tidak kritis	451 – 500	426 – 500	426 – 500

Sumber : Kemenhut, 2009

Sedangkan bobot dari parameter DAS untuk skoring kekritisian lahan di Kawasan Hutan ditunjukkan pada Tabel 2.2.

Tabel 2.2 Bobot dari parameter DAS untuk skoring kekritisian

Kawasan Hutan		Lindung	Budidaya	Luar kawasan
No.	Parameter DAS	Bobot	Bobot	Bobot
1.	Penutupa Lahan	50	30	50
2.	Kemiringan Lereng	20	20	20
3.	Tingkat Erosi	20	20	20
4.	Produktivitas Lahan	0	20	0
5.	Manajemen	10	10	10

Sumber : Kemenhut, 2009

Sebagai contoh daerah DAS di kawasan budidaya dengan karakteristik penutupan lahan pada kategori 2, Kemiringan lereng pada kategori 4, Tingkat Erosi pada kategori 2, Tingkat Produktivitas pada kategori 5 dan Manajemen pada kategori 5, penghitungan kekritisannya adalah sebagai berikut :

$$\begin{aligned}
 LK &= \text{Bobot} \times \text{Nilai Kategori} \\
 &= (2 \times 30) + (4 \times 20) + (2 \times 20) + (5 \times 20) + (5 \times 10) \\
 &= 60 + 80 + 40 + 100 + 50 \\
 &= 330 \text{ (Agak Kritis)}
 \end{aligned}$$

2.3 Parameter Karakteristik DAS

DAS memiliki karakteristik spesifik yang dicirikan oleh parameter-parameter yang berkaitan dengan keadaan morfometri, morfologi, tanah, geologi, vegetasi, tata guna (penggunaan) lahan, hidrologi, maupun yang berkaitan dengan manusia. Karakteristik DAS ini merupakan salah satu unsur utama dalam pengelolaan DAS seperti perencanaan serta monitoring dan evaluasi. Berdasarkan Permenhut Nomor P.32/Menhut-II/2009 tentang tata cara penyusunan rencana teknik rehabilitasi hutan dan lahan daerah aliran sungai, Parameter karakteristik DAS sebagai penentu tingkat kekritisian lahan yaitu meliputi tingkat penutupan lahan, kemiringan lereng, tingkat erosi, tingkat produktivitas serta manajemen lahan.

2.3.1 Penutupan Lahan

Parameter penutupan lahan dinilai berdasarkan prosentase penutupan tajuk pohon terhadap luas setiap *land system* dan diklasifikasikan menjadi lima kelas. Semakin baik tingkat penutupan lahan maka semakin lahan tersebut tidak mengalami kekritisian dan begitupula sebaliknya. Kategori penutupan lahan untuk masing-masing kelas ditunjukkan pada Tabel 2.3.

Tabel 2.3. Kategori dan Skoring Penutupan Lahan.

Kategori	Prosentase Penutupan Tajuk (%)	Skor
Sangat Baik	> 80	5
Baik	61-80	4
Sedang	41-80	3
Buruk	21-40	2
Sangat Buruk	< 20	1

Sumber : Kemenhut, 2009

2.3.2 Tingkat Kemiringan Lereng

Kemiringan lereng adalah perbandingan antara beda tinggi (jarak vertikal) suatu lahan dengan jarak mendatarnya. Besar kemiringan lereng dapat dinyatakan dengan beberapa satuan, diantaranya adalah dengan % (persen) dan o (derajat). Kategori kemiringan lereng untuk masing-masing kelas ditunjukkan pada Tabel 2.4.

Tabel 2.4. Kategori dan Skoring Kemiringan Lereng.

Kategori	Kemiringan Lereng (%)	Skor
Datar	< 8	5
Landai	8-15	4
Agak Curam	16-25	3
Curam	26-40	2
Sangat Curam	> 40	1

Sumber : Kemenhut, 2009

2.3.3 Tingkat Bahaya Erosi

Tingkat Bahaya Erosi (TBE) adalah perbandingan tingkat erosi di suatu satuan lahan dan kedalaman tanah efektif pada satuan lahan tersebut. Semakin Ringan tingkat bahaya erosi maka semakin lahan tersebut tidak mengalami kerusakan. Kategori tingkat erosi untuk masing-masing kelas ditunjukkan pada Tabel 2.5.

Tabel 2.5. Kategori dan Skoring Tingkat Erosi.

Kategori	Erosi (ton/ha/th)	Skor
Ringan	< 60	5
Sedang	60-180	4
Berat	180-480	3
Sangat Berat	> 480	2

Sumber : Kemenhut, 2009

2.3.4 Tingkat Produktivitas

Tingkat Produktivitas adalah rasio produktivitas lahan terhadap produksi komoditi umum optimal pada pengelolaan lahan secara tradisional. Semakin tinggi tingkat produktivitas lahan maka semakin lahan tersebut tidak mengalami kekritisian. Kategori tingkat produktivitas untuk masing-masing kelas ditunjukkan pada Tabel 2.6.

Tabel 2.6. Kategori dan Skoring Tingkat Produktivitas

Kategori	Tingkat Produktivitas (%)	Skor
Sangat tinggi	< 80	5
Tinggi	61-80	4
Sedang	41-60	3
Rendah	21-40	2
Sangat rendah	< 20	1

Sumber : Kemenhut, 2009

2.3.5 Manajemen

Manajemen merupakan salah satu kriteria yang dipergunakan untuk menilai lahan yang dinilai berdasarkan kelengkapan aspek pengelolaan

yang meliputi keberadaan tata batas kawasan, pengamanan dan pengawasan serta dilaksanakan atau tidaknya penyuluhan. Kategori tingkat manajemen untuk masing-masing kelas ditunjukkan pada Tabel 2.7.

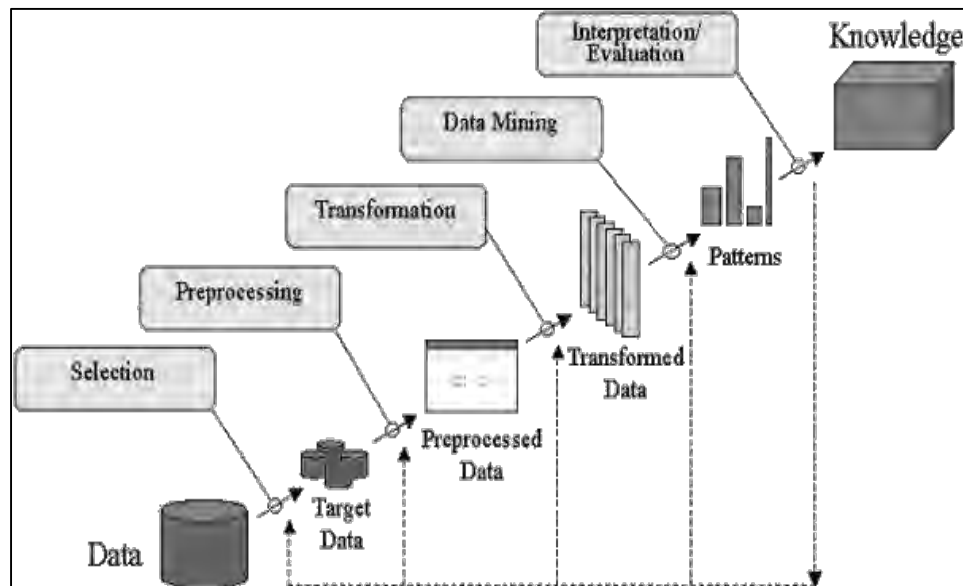
Tabel 2.7. Kategori dan Skoring Manajemen

Kategori	Besaran/Deskripsi	Skor
Baik	Lengkap	5
Sedang	Tidak Lengkap	3
Buruk	Tidak Ada	1

Sumber : Kemenhut, 2009

2.4 Data Mining

Data mining adalah metode pengolahan data untuk menemukan pola dan informasi yang tersembunyi dari sekumpulan data tersebut. Hasil dari pengolahan data dengan metode *data mining* ini dapat menghasilkan informasi yang digunakan untuk mengambil keputusan di masa depan. *Data mining* ini juga dikenal dengan istilah *pattern recognition*. Tahapan dalam data mining ditunjukkan dalam Gambar 2.2.



Sumber : gusconstan.com

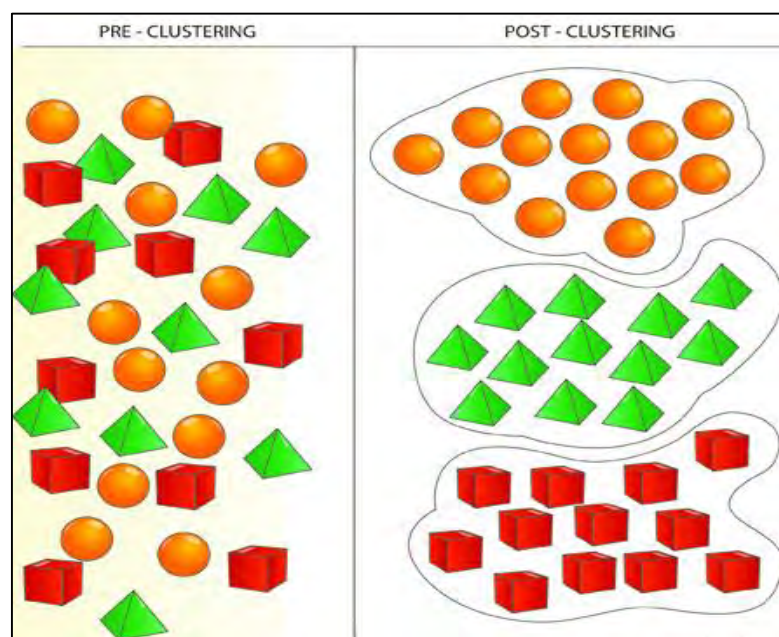
Gambar 2.2. Tahapan dalam Data Mining

Data mining merupakan metode pengolahan data berskala besar oleh karena itu data mining memiliki peranan penting dalam berbagai bidang seperti pemerintahan, industri, keuangan serta ilmu dan teknologi. Secara umum kajian data mining membahas metode-metode seperti prediksi, estimasi, clustering, klasifikasi, asosiasi, regresi, seleksi variable, dan market basket analisis (Santosa, 2007).

2.5 Clustering

Clustering merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan data yang lain. *Clustering* merupakan salah satu metode *data mining* yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*taining*) dan tanpa ada guru (*teacher*) serta tidak memerlukan target output.

Dalam data mining ada dua jenis metode clustering yang digunakan dalam pengelompokan data, yaitu *hierarchical clustering* dan *non-hierarchical clustering* (Santosa, 2007). Ilustrasi data clustering ditunjukkan pada Gambar 2.3.



Sumber :analyticstraining.com

Gambar 2.3. Ilustrasi data Clustering

2.5.1 Hierarical Clustering

Hierarical Clustering atau pengelompokan berbasis hierarki adalah suatu metode analisis kelompok yang berusaha untuk membangun sebuah hierarki cluster. Strategi pengelompokan berbasis hierarki umumnya ada 2(dua) jenis, yaitu aglomeratif dan divisif (Prasetyo, 2014).

Aglomeratif merupakan metode pengelompokan berbasis hierarki dengan pendekatan *bottom up*, yaitu proses pengelompokan dimulai dari masing-masing data sebagai sebuah cluster, kemudian secara rekursif mencari cluster terdekat sebagai pasangan untuk bergabung sebagai cluster yang lebih besar (Prasetyo, 2014). Proses tersebut diulang terus menerus sehingga tampak bergerak ke atas membentuk hierarki. Cara ini membutuhkan suatu parameter kedekatan cluster (*cluster proximity*). Divisif merupakan metode pengelompokan berbasis hierarki dengan pendekatan *top down*, yaitu pengelompokan dimulai dari suatu cluster yang berisi semua data, kemudian secara rekursif memecah cluster menjadi 2(dua) cluster sampai setiap cluster hanya berisi satu data tunggal (data itu sendiri). Untuk cara ini, yang dibutuhkan adalah keputusan cluster yang manakah yang akan dipecah pada setiap langkah dan bagaimana cara memecahnya (Prasetyo, 2014).

Pengelompokan berbasis hierarki sering ditampilkan dalam bentuk grafis menggunakan diagram mirip pohon (*tree*) yang disebut *dendogram*. Dendogram merupakan diagram yang menampilkan hubungan cluster dan subclusternya dalam urutan yang mana cluster yang digabung (*agglomerative view*) atau dipecah (*divisive view*).

2.5.2 Non Hierarical Clustering

Berbeda dengan clustering berbasis hierarki, *Non Hierarical Clustering* atau clustering non hierarki dimulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan (dua, tiga, dst). Setelah jumlah cluster ditentukan, maka proses cluster dilakukan dengan tanpa mengikuti proses hirarki yaitu dengan menggabungkan obyek ke dalam cluster-cluster tersebut.

Metode clustering non hierarki melakukan pengelompokan dengan terlebih dahulu memilih satu obyek dasar yang akan dijadikan nilai awal cluster, kemudian semua obyek yang ada didalam jarak terdekat dengan cluster ini akan bergabung lalu dipilih cluster kedua dan semua obyek yang mempunyai kemiripan dimasukkan dalam cluster ini. Demikian seterusnya hingga terbentuk beberapa cluster dengan keseluruhan obyek didalamnya.

2.6 Analisis Klaster

Analisis klaster digunakan untuk mengelompokkan data observasi yang hanya berdasarkan pada informasi yang ditemukan dalam data, di mana data tersebut harus menggambarkan observasi dan hubungannya. Oleh karena itu, tujuan dari analisis ini adalah observasi dalam satu kelompok mirip satu sama lain dan berbeda dari observasi dalam kelompok lain. Semakin besar kemiripan (homogenitas) dalam kelompok dan semakin besar perbedaan (heterogenitas) antar kelompok maka klastering akan lebih baik atau lebih berbeda (Tan *et al*, 2006).

Dalam analisis klaster, pengelompokan observasi ke dalam klaster dilakukan dengan menggunakan teknik-teknik yang berawal dari kemiripan antar semua pasangan observasi. Kemiripan ini didasarkan pada beberapa ukuran jarak. Metode lain dalam pengelompokan dapat menggunakan pilihan awal sebagai pusat klaster atau perbandingan di dalam dan antar variabilitas klaster. Selain itu, pengelompokan juga dapat menggunakan variabel klaster yang kemiripannya didasarkan pada matriks korelasi (Rencher, 2002).

Pada prinsipnya analisis klaster merupakan proses untuk mereduksi sejumlah objek yang besar menjadi lebih sedikit yang disebut klaster. Analisis klaster digunakan oleh peneliti yang belum mengetahui anggota dari suatu kelompok. Analisis klaster disebut juga *Q-analysis*, *classification analysis*, pengenalan pola (*pattern recognition*), analisis segmentasi (*numerical taxonomy*) (Handoko, 2012).

Berdasarkan paparan tersebut, terdapat dua langkah utama dalam analisis klaster yaitu memilih ukuran kemiripan dan memilih algoritma

dalam pembentukan klaster (Handoko, 2012).

1. Tujuan Analisis Klaster

Setelah mengelompokkan n buah objek pengamatan kedalam m kelompok berdasarkan p variat dapat diketahui bahwa tujuan utama dari pengklasteran objek adalah untuk mendapatkan kelompok objek yang memiliki nilai relatif sama. Sehingga kelak dalam interpretasi, objek-objek yang berada pada satu klaster memiliki peluang yang cukup tinggi akan muncul bersamaan pada satu individu.

2. Konsep Dasar dalam Analisis Klaster

Analisis klaster merupakan suatu kelas teknik, dipergunakan untuk mengklasifikasi objek atau kasus ke dalam kelompok yang relatif homogen, yang disebut klaster. Objek dalam setiap kelompok cenderung mirip satu sama lain dan berbeda jauh (tidak sama) dengan objek dari klaster lainnya (Supranto, 2004).

2.7 Algoritma *K-Modes Clustering*

Metode yang banyak digunakan dalam pengelompokan non hierarki adalah *K-Means Clustering*. K-means clustering merupakan algoritma pengelompokan iteratif yang melakukan partisi set data kedalam sejumlah K cluster yang sudah ditetapkan diawal. Algoritma K-means sederhana untuk diimplementasikan dan dijalankan, relatif cepat, mudah beradaptasi serta umum penggunaanya dalam praktek. Secara historis K-means menjadi salah satu algoritma yang paling penting dalam bidang data mining (Wu dan Kumar, 2009).

Algoritma *K-Modes Clustering* adalah modifikasi dari Algoritma K-means Clustering yang digunakan untuk fitur kategorikal (nominal atau ordinal). K-means hanya dapat bekerja dengan baik untuk set data yang tipe data fiturnya numeric (interval atau rasio) (Prasetyo, 2014). Contoh data set dengan fitur kategorikal ditunjukkan pada Tabel 2.8.

Tabel 2.8 Contoh data set kategorikal 3 dimensi

Data Ke -i	Fitur1	Fitur2	Fitur3
1	1	3	2
2	2	3	2
3	2	1	2

Keterangan Kategori:

1 = Buruk, 2 = Cukup, 3 = Baik

Untuk menyelesaikan masalah tersebut. K-modes melakukan modifikasi pada K-means sebagai berikut (Prasetyo, 2014):

1. Menggunakan ukuran pencocokan ketidakmiripan sederhana pada fitur data bertipe kategorikal.
2. Mengganti mean cluster dengan modus (nilai yang paling sering muncul).
3. Menggunakan metode berbasis frekuensi untuk mencari modus dari sekumpulan nilai.

Andaikan x dan y adalah dua data dengan fitur bertipe kategorikal. Ukuran ketidakmiripan diantara x dan y dapat diukur dengan jumlah ketidakcocokan nilai dari fitur yang berkorespondensi dari dua data. Semakin kecil ketidakcocokan, maka semakin mirip dua data tersebut, Metric seperti ini sering disebut dengan pencocokan sederhana (*simple matching*). Formula penghitungan jarak dalam K-Modes ditunjukkan seperti pada persamaan 1.

$$d(x, y) = \sum_{j=1}^r \epsilon(x_j - c_j) \quad \dots(1)$$

Dimana :

$d(x, c)$ = Jarak data x ke c

x_j = Nilai fitur ke- j dari x

c_j = Nilai fitur ke- j dari c

r = Jumlah fitur dalam vector

dengan :

$$\epsilon(x_j, c_j) = \begin{pmatrix} 0 & (x_j = c_j) \\ 1 & (x_j \neq c_j) \end{pmatrix}$$

langkah langkah melakukan Algoritma *K-Modess Clustering* adalah sebagai berikut (Prasetyo, 2014):

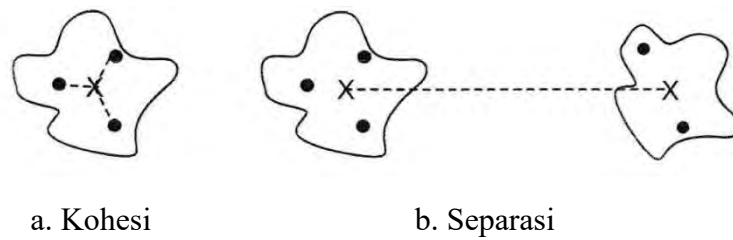
1. Tentukan nilai k sebagai inisialisasi centroid awal untuk setiap cluster.
2. Alokasikan semua data ke *centroid* terdekat menggunakan persamaan 1.
3. Perbarui modus (sebagai centroid baru) dari setiap cluster dengan nilai kategori yang sering muncul pada setiap cluster .
4. Ulangi langkah 2 dan 3 selama memenuhi syarat, yaitu (a) tidak ada data yang berpindah cluster atau (b) posisi *centroid* tidak berubah

2.8 Validitas Cluster

Validitas merupakan proses untuk menilai hasil algoritma cluster yang terbentuk, untuk menjamin bahwa solusi cluster yang dihasilkan dapat menggambarkan populasi yang sebenarnya (Yuhefizar, 2013). Ukuran validitas cluster relatif adalah ukuran evaluasi *unsupervised* ataupun *supervised* yang digunakan untuk tujuan pembandingan. Misalnya dua clustering dengan K-means dapat dibandingkan menggunakan nilai validitas cluster tersebut dengan melakukan evaluasi validitas cluster. Basis metode yang berbeda juga memberikan cara evaluasi yang berbeda. Metode berbasis non hierarki seperti K-Means, metode fuzzy, maupun pengelompokkan pada data kategorikal juga memerlukan cara evaluasi yang berbeda (Prasetyo, 2014).

2.7.1 Validitas Internal

Banyak metrik internal yang mengukur validitas cluster pada metode pengelompokkan berbasis non hierarki didasarkan pada nilai kohesi dan separasi. Kohesi dalam pengelompokkan berbasis non hierarki didefinisikan sebagai jumlah dari kedekaan data terhadap centroid dari cluster yang diikutinya. Sedangkan separasi diantara dua cluster dapat diukur dengan kedekatan dua prototipe (centroid) cluster (Prasetyo, 2014). Gambaran tentang kohesi dan separasi digambarkan pada gambar 2.4, centroid dalam Gambar 2.4 disimbolkan dengan tanda “X” .



Gambar 2.4. Konsep Kohesi dan Separasi

Beberapa metode validitas cluster secara internal antara lain *Dunn Index*, *Davies Bouldin Index* dan *Silhouette Index*.

2.7.2 Validitas Eksternal

Jika evaluasi cluster dilakukan tanpa informasi eksternal (misal label kelas) maka evaluasi tersebut disebut juga dengan *unsupervised*, tapi jika ada informasi eksternal yang dimiliki setiap data yang dikelompokkan maka disebut dengan *supervised*. Validitas yang dilakukan biasanya adalah mengukur tingkat hubungan antara label cluster dengan label kelas. Alasan melakukan evaluasi cluster menggunakan informasi eksternal adalah perbandingan teknik clustering dengan kebenaran dasar atau sejauh mana proses panduan klasifikasi dapat secara otomatis dihasilkan oleh analisis cluster (Tan et al, 2006).

Ada 2(dua) teknik yang dapat digunakan yaitu berorientasi klasifikasi (*classification oriented*) dan berorientasi kemiripan (*similarity oriented*). Metode berorientasi klasifikasi mengukur sejauh mana cluster berisi objek dari kelas tunggal. Sementara metode berorientasi kemiripan mengukur sejauh mana dua objek yang berada dalam kelas yang sama berada dalam cluster yang sama, begitu pula sebaliknya.

2.7.3 Davies Bouldin Index

Davies-Bouldin Index (DBI) adalah salah satu metode validitas internal dalam melakukan evaluasi terhadap suatu cluster. Metrik *DBI* diperkenalkan oleh David L Davies dan Donald W. Bouldin (1979), Validitas Internal yang dilakukan DBI adalah seberapa baik cluster sudah dilakukan dengan menghitung kuantitas dan fitur turunan dari set data (Prasetyo, 2014).

Sum of square within cluster (SSW) sebagai metrik kohesi dalam sebuah cluster ke- i diformulasikan pada persamaan (2).

$$SSW_i = \frac{1}{m_i} \sum_{j=1}^{m_i} d(x_j, c_i) \quad \dots(2)$$

Dimana :

SSW = Sum of square within cluster

m_i = Jumlah data dalam cluster i

$d(x, c)$ = Jarak data x ke centroid c

Nilai $d()$ dalam persamaan (2) bisa menggunakan formula ketidakmiripan (jarak) yang digunakan ketika proses pengelompokkannya sehingga validasi yang diberikan juga mempunyai maksud yang sama terhadap proses pengelompokkannya.

Sementara metrik untuk separasi antara dua cluster digunakan formula Sum of square between cluster (SSB) dengan mengukur jarak antar centroid c_i dan c_j seperti pada persamaan (3).

$$SSB_{i,j} = d(c_i, c_j) \quad \dots(3)$$

Dimana :

SSB = Sum of square between cluster

$d(c_i, c_j)$ = Jarak centroid c_i dengan centroid c_j

Langkah selanjutnya adalah menghitung nilai R_{ij} . R_{ij} adalah ukuran rasio seberapa baik nilai perbandingan antara cluster ke- i dan cluster ke- j . Nilainya didapatkan dari komponen kohesi dan separasi. Cluster yang baik adalah yang mempunyai kohesi yang sekecil mungkin dan separasi yang sebesar mungkin. R_{ij} di formulasikan dalam persamaan (4).

$$R_{ij} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad \dots(4)$$

Dimana :

R_{ij} = Nilai Perbandingan cluster i dengan cluster j

Setelah kita mendapatkan nilai SSW , SSB dan R_{ij} . Nilai *Davies-Bouldin Index* (DBI) dapat dihitung menggunakan formula dalam persamaan (5).

$$DBI = \frac{1}{K} \sum_{j=1}^K \max(R_{i,j}) \quad \dots(5)$$

Dimana :

K = Jumlah cluster yang digunakan

Dari formula diatas dapat kita amati bahwa semakin rendah nilai DBI maka semakin baik cluster yang didapatkan, artinya kemiripan antar data dalam satu cluster akan semakin mirip. DBI banyak digunakan untuk membantu clustering berbasis non hierarki seperti K-Means atau K-Modes untuk menentukan berapa jumlah cluster yang tepat untuk digunakan.

2.9 Sistem Informasi Geografis (SIG)

Sistem informasi geografis (SIG) adalah suatu sistem berbasis komputer yang mempunyai kemampuan untuk menyimpan, menganalisis serta menampilkan data geografis (Chang, 2002). SIG dapat dibagi menjadi 4 (empat) komponen (Chang, 2002) yaitu :

1. Sistem Komputer, komponen ini mencakup perangkat keras dan sistem operasi untuk menjalankan SIG.
2. Perangkat lunak SIG, komponen ini mencakup program aplikasi SIG untuk menjalankan perangkat keras.
3. *Brainware*, komponen ini menunjuk pada tujuan dan sasaran serta alasan dalam penggunaan SIG.
4. Infrastruktur, Komponen ini menunjuk pada organisasi, administratif dan lingkungan untuk operasi SIG.

Selanjutnya SIG akan selalu diasosiasikan dengan sistem yang berbasis komputer, walaupun pada dasarnya SIG dapat dikerjakan secara manual, SIG yang berbasis komputer akan sangat membantu ketika data geografis merupakan data yang besar (dalam jumlah dan ukuran) dan terdiri dari banyak tema yang saling berkaitan.

Data yang akan diolah dalam SIG merupakan data spasial yaitu sebuah data yang berorientasi geografis dan merupakan lokasi yang memiliki sistem koordinat tertentu, sebagai dasar referensinya.

2.9.1 Data dan Operasi dalam SIG

Dalam kerangka kerja SIG bentuk dan struktur data dibagi menjadi 2 (dua) kategori (Chang, 2002) yaitu :

1. Data Spasial, data spasial dapat dibedakan menjadi dua yaitu data vektor dan data raster. Data vektor menggunakan koordinat untuk membentuk fitur-fitur spasial seperti titik, garis dan bidang. Sedangkan data raster menggunakan grid untuk merepresentasikan variasi spasial.
2. Data atribut, data atribut mendeskripsikan karakteristik dari fitur-fitur spasial .

Sedangkan Operasi dalam SIG dapat digolongkan menjadi enam kelompok (Chang, 2002) yaitu :

1. Input Data, dalam operasi ini dilakukan pemasukan data.
2. Manajemen dan atribut, dalam operasi ini dilakukan verifikasi data dan manajemen basis data.
3. Menampilkan data menggunakan peta. Tabel dan grafik
4. Eksplorasi data, eksplorasi data merupakan *query* dan analisis berpusatkan pada data.
5. Analisis data, Analisis data vektor meliputi *buffering*, *overlay*, *distance measure* dan manipulasi peta. Analisis data raster meliputi *local neighborhood*, *zonal* dan *global*.
6. Pemodelan SIG, pemodelan SIG menunjuk pada penggunaan SIG dalam pembuatan model analisis.

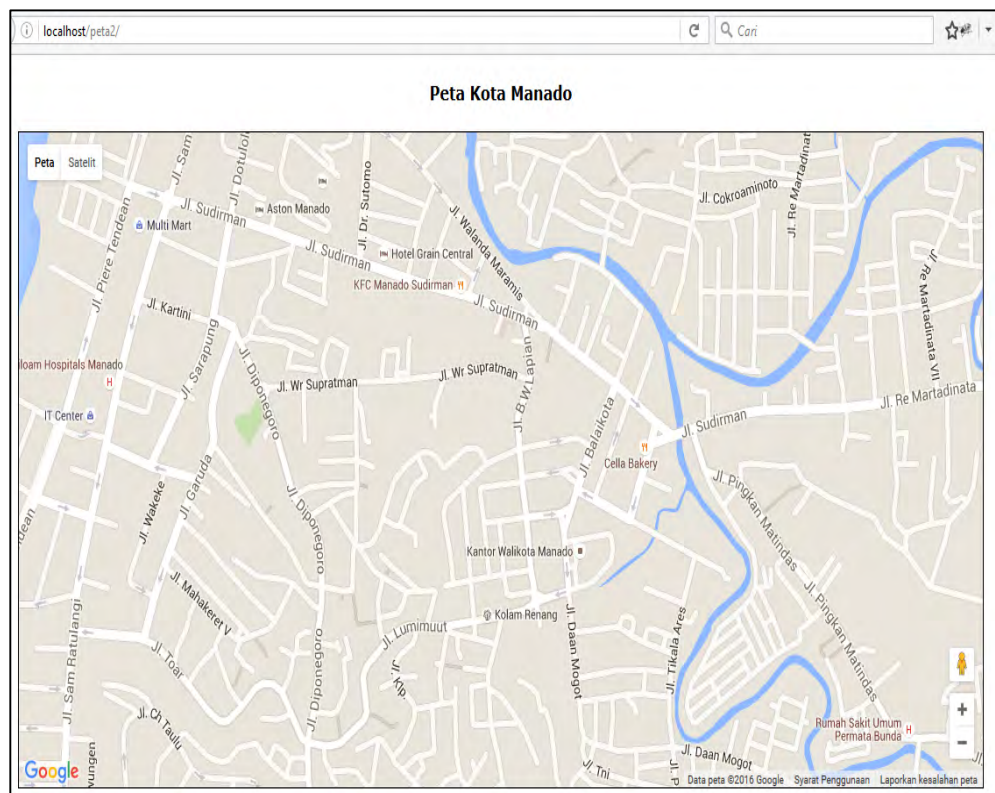
2.9.2 Google Map API

Perkembangan SIG dengan segala kelebihananya sangatlah pesat. Pemanfaatan SIG semakin meluas meliputi berbagai bidang seperti sumber daya alam, ekonomi, kesehatan, pertanian dan sebagainya begitu pula dengan aplikasinya. salah satu aplikasi SIG yang sangat populer dan banyak digunakan saat ini adalah Google Map.

Google Maps adalah layanan aplikasi dan teknologi peta berbasis web yang disediakan Google secara gratis (bukan untuk kepentingan

komersial). Google menambahkan fasilitas peta yang dinamis dengan menambahkan fitur Pemograman sendiri dengan Google Maps API (*Application Programming Interface*). Google Maps API sendiri merupakan library dari bahasa pemograman JavaScript.

Dengan menggunakan Google Maps API Programmer dapat menghemat waktu dan biaya untuk membangun aplikasi peta digital yang handal, sehingga Programmer dapat fokus hanya pada data yang akan disajikan karena peta sudah disediakan oleh Google. Contoh peta berbasis Google Map API di tunjukan pada Gambar 2.5



Gambar 2.5. Contoh peta kota Manado berbasis Google Map API

Google Map API dapat digabungkan dengan beberapa bahasa pemograman lain seperti PHP, HTML, AJAX serta DBMS seperti MySQL, Oracle maupun yang lain. Sehingga memungkinkan kita untuk melakukan operasi dalam SIG seperti Input Data, manajemen atribut, analisis data serta melakukan pemodelan SIG.

2.10 Penelitian Terkait

Beberapa penelitian mengenai Data mining menggunakan algoritma K-modes clustering sebelumnya telah dilakukan. Salah satunya oleh Handayani dengan judul “ *Analisis Clustering menggunakan Algoritma K-Modes*”, dalam penelitian ini diperoleh analisa bahwa algoritma K-means hanya terbatas untuk pengelompokan pada data numerik, sedangkan untuk menangani masalah data kategorik diperlukan varian algoritma K-means yaitu K-modes, dari penelitian ini juga didapatkan bahwa pemilihan *centroid* awal menggunakan metode frequency based memiliki tingkat pengelompokan yang lebih baik dibandingkan inisialisasi *centorid* secara random (Handayani, 2010).

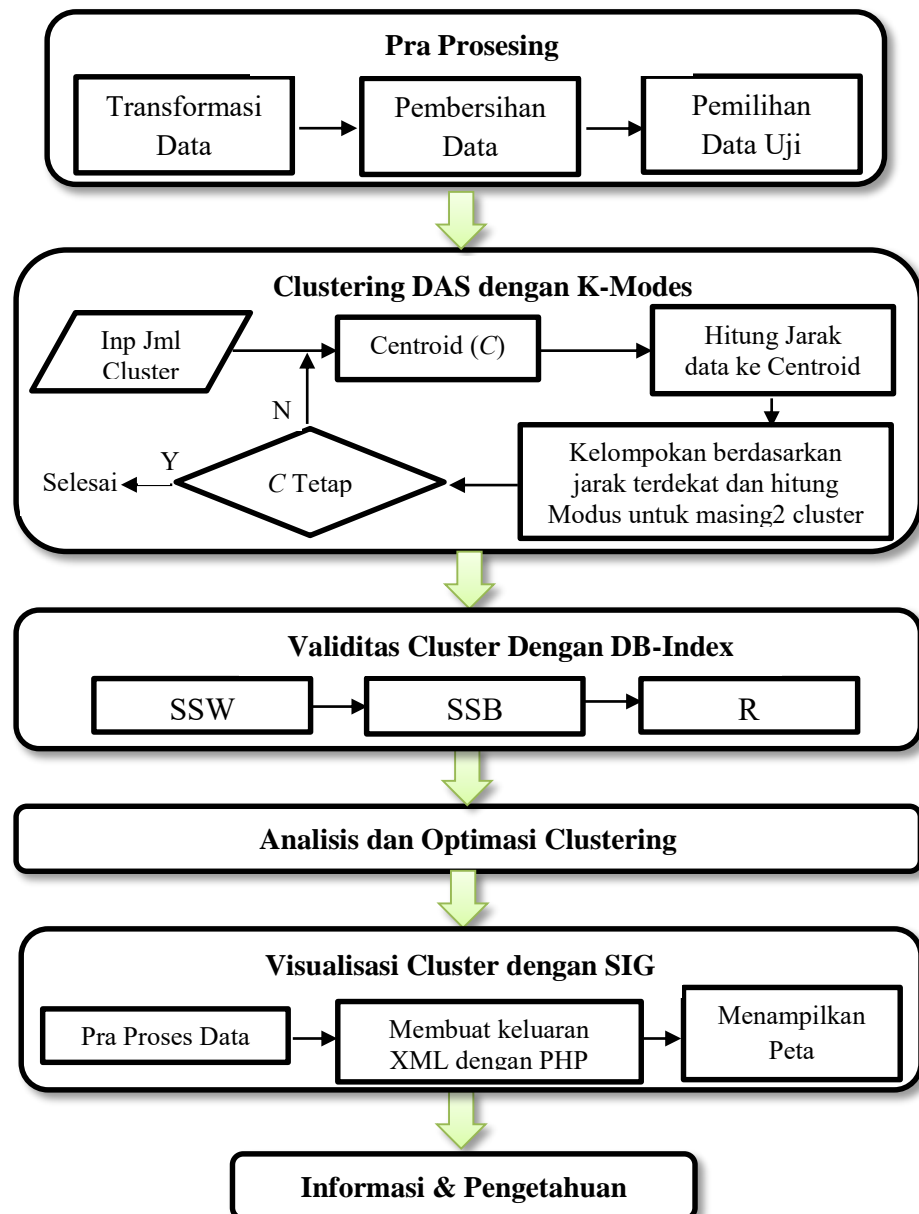
Firli,dkk melakukan Penelitian mengenai optimasi jumlah Cluster menggunakan Davies-Bouldin Index untuk pengelompokan kecamatan dengan judul “*Optimalisasi Pengelompokkan Kecamatan Berdasarkan Indikator Pendidikan Menggunakan Metode Clustering dan Davies Bouldin Index*” Hasil dari penelitian ini didapatkan jumlah empat cluster yang merupakan jumlah cluster optimal untuk pengelompokkan kecamatan berdasarkan indikator pendidikan (Firli dkk, 2014).

Penelitian tentang Sistem Informasi Geografis untuk visualisasi hasil clustering sebelumnya pernah dilakukan oleh Handoko dengan judul “*Sistem Informasi Geografis Berbasis Web untuk Pemetaan Sebaran Alumni menggunakan Metode K-Means*”. Dari penelitian ini didapatkan pola sebaran alumni yang sesuai, kurang sesuai dan tidak sesuai antara bidang kerja dan kompetensi pendidikan (Handoko, 2012). Sedangkan dari penelitian harianja dengan judul “ *Visualisasi K-Means Clustering pada Data Potensi Pertanian Desa di Bogor Menggunakan MapServer*” diperoleh sebaran kelompok desa dengan tingkat potensi pertanian antara kecil, sedang dan besar (Harianja, 2008).

BAB III

METODOLOGI

Penelitian ini dilakukan berdasarkan beberapa tahapan yang bertujuan memperlihatkan bagaimana sebuah model *clustering* bisa memberikan informasi dan pengetahuan tentang karakteristik kelompok DAS berdasarkan parameter yang ada. Tahapan dalam penelitian ini ditunjukkan pada Gambar 3.1.



Gambar 3.1. Diagram alur penelitian

3.1 Pra Processing Data

Tahapan ini meliputi proses transformasi data, pembersihan data dan pemilihan data uji. Proses transformasi data akan mengekspor data awal yang diperoleh dari Balai Pegelolaan Daerah Aliran Sungai (BPDAS) Tondano dalam format MS Excel (.xlsx) ke database MySQL. Kemudian dilakukan proses pembersihan data, atribut yang tidak digunakan akan dibuang sebagian, field data yang kosong serta isi field tidak sesuai dengan kategori yang ada sehingga akan mengurangi jumlah data asal.

Setelah proses pembersihan data kemudian dilakukan pemilihan data uji. Proses ini dilakukan untuk memilih data yang akan diujikan. Dari pra processing data ini akan menghasilkan data yang siap untuk di lakukan *clustering* di tahap selanjutnya.

3.2 Clustering DAS dengan Algoritma *K-Modes Clustering*

Pada proses ini akan dilakukan clustering menggunakan algoritma k-modes clustering, misalkan kita mempunyai contoh data DAS dengan parameter karakteristik seperti ditunjukkan pada Tabel 3.1.

Tabel 3.1 Dataset 10 data DAS

Data Ke-	Kategori				
	PL	Lereng	Erosi	Prod	Man
1	3	5	2	3	3
2	5	3	4	5	5
3	1	5	2	3	5
4	3	3	2	3	5
5	2	2	3	2	3
6	3	3	2	3	5
7	1	5	3	3	3
8	4	2	4	4	5
9	1	2	2	3	3
10	2	2	2	2	3

Data akan di bagi menjadi 3 cluster, Tahapan clustering dengan algoritma K-modes clustering adalah sebagai berikut :

1. Tentukan nilai k sebagai inisialisasi centroid (modus) awal untuk setiap cluster, misalnya dipilih data ke-2, ke-5 dan ke-9, seperti ditunjukkan pada tabel 3.2.

Tabel 3.2 Centroid awal pada contoh data DAS

Centroid	Kategori				
	PL	Lereng	Erosi	Prod	Man
1	5	3	4	5	5
2	2	2	3	2	3
3	1	2	2	3	3

2. Hitung jarak semua data ke *centroid* menggunakan persamaan 1, berikut perhitungan jarak ke setiap centroid pada data ke-1.

$$\begin{aligned}
 d(x_1, c_1) &= \epsilon(x_{11}, c_{11}) + \epsilon(x_{12}, c_{12}) + \epsilon(x_{13}, c_{13}) + \\
 &\quad \epsilon(x_{14}, c_{14}) + \epsilon(x_{15}, c_{15}) \\
 &= \epsilon(3,5) + \epsilon(5,3) + \epsilon(2,4) + \epsilon(3,5) + \epsilon(3,5) \\
 &= 1+1+1+1+1 = \mathbf{5}
 \end{aligned}$$

$$\begin{aligned}
 d(x_1, c_2) &= \epsilon(x_{11}, c_{21}) + \epsilon(x_{12}, c_{22}) + \epsilon(x_{13}, c_{23}) + \\
 &\quad \epsilon(x_{14}, c_{24}) + \epsilon(x_{15}, c_{25}) \\
 &= \epsilon(3,2) + \epsilon(5,2) + \epsilon(2,3) + \epsilon(3,2) + \epsilon(3,3) \\
 &= 1+1+1+1+0 = \mathbf{4}
 \end{aligned}$$

$$\begin{aligned}
 d(x_1, c_3) &= \epsilon(x_{11}, c_{31}) + \epsilon(x_{12}, c_{32}) + \epsilon(x_{13}, c_{33}) + \\
 &\quad \epsilon(x_{14}, c_{34}) + \epsilon(x_{15}, c_{35}) \\
 &= \epsilon(3,1) + \epsilon(5,2) + \epsilon(2,2) + \epsilon(3,3) + \epsilon(3,3) \\
 &= 1+1+0+0+0 = \mathbf{2}
 \end{aligned}$$

Selanjutnya dilakukan perhitungan jarak untuk data ke-2 sampai data ke-10 ke setiap centroid, sehingga menghasilkan matrik jarak dan cluster yang diikuti pada iterasi ke-1, seperti ditunjukkan pada Tabel 3.3.

Tabel 3.3 Matrik jarak dan cluster yang diikuti pada iterasi ke-1 data DAS.

Data Ke-	Kategori					Jarak ke centroid			Ter dek at	Cluster diikuti
	PL	Ler eng	Ero si	Prod	Man	1	2	3		
1	3	5	2	3	3	5	4	2	2	3
2	5	3	4	5	5	0	5	5	0	1
3	1	5	2	3	5	4	5	2	2	3
4	3	3	2	3	5	3	5	3	3	1
5	2	2	3	2	3	5	0	3	0	2
6	3	3	2	3	5	3	5	3	3	1
7	1	5	3	3	3	5	3	2	2	3
8	4	2	4	4	5	3	4	4	3	1
9	1	2	2	3	3	5	3	0	0	3
10	2	2	2	2	3	5	1	2	1	2

- Perbarui modus (sebagai centroid baru) dari setiap cluster berdasarkan data yang bergabung pada setiap clusternya. Untuk cluster 1 ada 4 data yang tergabung didalamnya, untuk cluster 2 ada 2 data, sedangkan untuk cluster 3 ada 4 data yang tergabung. seperti ditunjukkan pada Tabel 3.4 , 3.5 dan Tabel 3.6

Tabel 3.4 Data yang tergabung pada cluster 1 data DAS

Data Ke-	Kategori				
	PL	Lereng	Erosi	Prod	Man
2	5	3	4	5	5
4	3	3	2	3	5
6	3	3	2	3	5
8	4	2	4	4	5

Tabel 3.5 Data yang tergabung pada cluster 2 data DAS

Data Ke-	Kategori				
	PL	Lereng	Erosi	Prod	Man
5	2	2	3	2	3
10	2	2	2	2	3

Tabel 3.6 Data yang tergabung pada cluster 3 data DAS

Data Ke-	Kategori				
	PL	Lereng	Erosi	Prod	Man
1	3	5	2	3	3
3	1	5	2	3	5
7	1	5	3	3	3
9	1	2	2	3	3

Modus yang didapatkan dari 3 cluster tersebut, seperti ditunjukkan pada Tabel 3.7 .

Tabel 3.7 centroid baru yang didapat.

Centroid baru	Kategori				
	PL	Lereng	Erosi	Prod	Man
1	3	3	2	3	5
2	2	2	3	2	3
3	1	5	2	3	3

4. Lanjutkan Iterasi ke-2 dengan mengulangi langkah 2 dan 3 menggunakan centroid yang baru sampai memenuhi syarat, yaitu (a) tidak ada data yang berpindah cluster atau (b) posisi *centroid* tidak berubah.

3.3 Validitas Cluster dengan Metode *Davies Bouldin Index*

Setelah proses clustering menghasilkan beberapa cluster, poses selanjutnya dalam penelitian ini adalah menghitung nilai validitas cluster menggunakan *Davies Bouldin Index* (DBI). misalkan kita mempunyai contoh hasil clustering data DAS seperti ditunjukkan pada Tabel 3.8 dan centroid akhir seperti ditunjukkan pada Tabel 3.9

Tabel 3.8 Contoh Hasil Clustering data DAS

Data Ke-	Kategori					Cluster akhir
	PL	Lereng	Erosi	Prod	Man	
1	3	5	2	3	3	3
2	5	3	4	5	5	1
3	1	5	2	3	5	3
4	3	3	2	3	5	1
5	2	2	3	2	3	2
6	3	3	2	3	5	1
7	1	5	3	3	3	3
8	4	2	4	4	5	1
9	1	2	2	3	3	3
10	2	2	2	2	3	2

Tabel 3.9 Contoh centroid akhir Clustering data DAS

Centroid akhir	Kategori				
	PL	Lereng	Erosi	Prod	Man
1	3	3	2	3	5
2	2	2	3	2	3
3	1	5	2	3	3

Tahapan menghitung Nilai DBI adalah sebagai berikut :

1. Kelompokkan data berdasarkan clusternya, kemudian dihitung jarak setiap data dengan centroid akhir masing-masing (menggunakan formula ketidakmiripan sesuai persamaan 1) dan hitung rata-ratanya

untuk menjadi Sum of square within cluster (SSW). Berikut perhitungan jarak data dalam cluster 1 ke centroid akhir-nya.

$$\begin{aligned}
 d(x_2, c_1) &= \epsilon(x_{21}, c_{11}) + \epsilon(x_{22}, c_{12}) + \epsilon(x_{23}, c_{13}) + \\
 &\quad \epsilon(x_{24}, c_{14}) + \epsilon(x_{25}, c_{15}) \\
 &= \epsilon(5,3) + \epsilon(3,3) + \epsilon(4,2) + \epsilon(5,3) + \epsilon(5,5) \\
 &= 1+0+1+1+0 = \mathbf{3}
 \end{aligned}$$

$$\begin{aligned}
 d(x_4, c_1) &= \epsilon(x_{41}, c_{11}) + \epsilon(x_{42}, c_{12}) + \epsilon(x_{43}, c_{13}) + \\
 &\quad \epsilon(x_{44}, c_{14}) + \epsilon(x_{45}, c_{15}) \\
 &= \epsilon(3,3) + \epsilon(3,3) + \epsilon(2,2) + \epsilon(3,3) + \epsilon(5,5) \\
 &= 0+0+0+0+0 = \mathbf{0}
 \end{aligned}$$

$$\begin{aligned}
 d(x_6, c_1) &= \epsilon(x_{61}, c_{11}) + \epsilon(x_{62}, c_{12}) + \epsilon(x_{63}, c_{13}) + \\
 &\quad \epsilon(x_{64}, c_{14}) + \epsilon(x_{65}, c_{15}) \\
 &= \epsilon(3,3) + \epsilon(3,3) + \epsilon(2,2) + \epsilon(3,3) + \epsilon(5,5) \\
 &= 0+0+0+0+0 = \mathbf{0}
 \end{aligned}$$

$$\begin{aligned}
 d(x_8, c_1) &= \epsilon(x_{81}, c_{11}) + \epsilon(x_{82}, c_{12}) + \epsilon(x_{83}, c_{13}) + \\
 &\quad \epsilon(x_{84}, c_{14}) + \epsilon(x_{85}, c_{15}) \\
 &= \epsilon(4,3) + \epsilon(2,3) + \epsilon(4,2) + \epsilon(4,3) + \epsilon(5,5) \\
 &= 1+1+1+1+0 = \mathbf{4}
 \end{aligned}$$

Setelah jarak data dengan centroid untuk cluster 1 didapatkan, perhitungan nilai SSW untuk cluster 1 dapat menggunakan persamaan 2, sebagai berikut :

$$SSW_1 = \frac{1}{4} (d(x_2, c_1) + d(x_4, c_1) + d(x_6, c_1) + d(x_8, c_1))$$

$$SSW_1 = \frac{1}{4} (3 + 0 + 0 + 4) = 1.75$$

Nilai SSW untuk cluster 2 dan 3 dihitung menggunakan cara yang sama, sehingga menghasilkan matrik nilai SSW yang ditunjukkan pada Tabel 3.10

Tabel 3.10 Matrik nilai SSW contoh clustering data DAS

Data Ke-	Cluster yg diikuti						Jarak ke centroid akhir	SSW
		PL	Lere ng	Ero si	Prod	Man		
2	1	3	3	2	3	5	3	1.75
4	1						0	
6	1						0	
8	1						4	
5	2	2	2	3	2	3	0	0.5
10	2						1	
1	3	1	5	2	3	3	1	1
3	3						1	
7	3						1	
9	3						1	

2. Nilai SSB didapatkan menggunakan formula persamaan 2, dengan penghitungan jarak (menggunakan formula ketidakmiripan di persamaan 1) antar centroid, perhitungan SSB pasangan diantara 3 cluster tersebut sebagai berikut :

$$\begin{aligned}
 SSB_{12} &= d(c_1, c_2) \\
 &= \epsilon(c_{11}, c_{21}) + \epsilon(c_{12}, c_{22}) + \epsilon(c_{13}, c_{23}) + \epsilon(c_{14}, c_{24}) + \epsilon(c_{15}, c_{25}) \\
 &= \epsilon(3,2) + \epsilon(3,2) + \epsilon(2,3) + \epsilon(3,2) + \epsilon(5,3) \\
 &= 1+1+1+1+1 = \mathbf{5}
 \end{aligned}$$

$$\begin{aligned}
 SSB_{13} &= d(c_1, c_3) \\
 &= \epsilon(c_{11}, c_{31}) + \epsilon(c_{12}, c_{32}) + \epsilon(c_{13}, c_{33}) + \epsilon(c_{14}, c_{34}) + \epsilon(c_{15}, c_{35}) \\
 &= \epsilon(3,1) + \epsilon(3,5) + \epsilon(2,2) + \epsilon(3,3) + \epsilon(5,3) \\
 &= 1+1+0+0+1 = \mathbf{3}
 \end{aligned}$$

$$\begin{aligned}
SSB_{23} &= d(c_2, c_3) \\
&= \epsilon(c_{21}, c_{31}) + \epsilon(c_{22}, c_{32}) + \epsilon(c_{23}, c_{33}) + \epsilon(c_{24}, c_{34}) + \\
&\quad \epsilon(c_{25}, c_{35}) \\
&= \epsilon(2,1) + \epsilon(2,5) + \epsilon(3,2) + \epsilon(2,3) + \epsilon(3,3) \\
&= 1+1+1+1+0 = 4
\end{aligned}$$

Sehingga menghasilkan matrik nilai SSB yang ditunjukkan pada Tabel 3.11

Tabel 3.11 Matrik nilai SSB contoh clustering data DAS

		Data Centroid Ke- <i>i</i>		
SSB		1	2	3
Data Centroid Ke- <i>i</i>	1	0	5	3
	2	5	0	4
	3	3	4	0

3. Setelah didapatkan nilai SSW dan SSB, langkah selajutnya adalah menghitung nilai R_{ij} . R_{ij} adalah ukuran rasio seberapa baik nilai perbandingan antara cluster ke-*i* dan cluster ke-*j*. Nilainya didapatkan dari komponen kohesi dan separasi menggunakan persamaan 4. Perhitungan nilai R_{ij} pasangan diantara 2 cluster tersebut sebagai berikut :

$$R_{12} = \frac{SSW_1 + SSW_2}{SSB_{12}} = \frac{1.75 + 0.5}{5} = \frac{3.42}{5} = 0.45$$

$$R_{13} = \frac{SSW_1 + SSW_3}{SSB_{13}} = \frac{1.75 + 1}{3} = \frac{2.75}{3} = 0.917$$

$$R_{23} = \frac{SSW_2 + SSW_3}{SSB_{23}} = \frac{0.5 + 1}{4} = \frac{1.5}{4} = 0.375$$

4. Langkah terakhir adalah menghitung nilai DBI, Nilai DBI dapat dihitung menggunakan formula dalam persamaan (5). Dimana R_{max} pada masing cluster akan dirata-rata yang merupakan nilai DBI akhir. Perhitungan nilai R_{max} dan DBI hasil clustering ditunjukkan pada Tabel 3.12

Tabel 3.12 Nilai R_{max} dan DBI contoh clustering data DAS

R		Data Cluster Ke- i			R_{max}	DBI
		1	2	3		
Data Cluster	1	0	0.45	0.917	0.917	0.761
Ke- i	2	0.45	0	0.375	0.45	
	3	0.917	0.375	0	0.917	

Terlihat bahwa nilai DBI yang didapatkan adalah 0.761

3.4 Analisis dan Optimasi Clustering

Analisis dan optimasi dilakukan untuk mengetahui berapa jumlah cluster yang menghasilkan nilai DBI terendah, artinya jumlah cluster dengan tingkat kemiripan data yang paling mirip serta mendapatkan profiling karakteristik DAS dari masing-masing cluster maupun analisis dari sebaran DAS pada masing-masing cluster. Dalam penelitian ini percobaan akan dilakukan menggunakan 2 sampai 10 cluster dengan nilai *centroid* awal secara random sebanyak 10 kali.

Rekapitulasi hasil percobaan menggunakan 2 sampai 10 cluster dengan nilai *centroid* awal secara random sebanyak 10 kali Akan dianalisa sehingga didapatkan berapa jumlah cluster yang menghasilkan rata-rata nilai DBI terendah yang kemudian akan dijadikan jumlah cluster dalam clustering data DAS. Setelah cluster terbentuk, profil data pada setiap cluster akan membentuk segmen DAS berdasarkan karakteristiknya

3.5 Visualisasi Cluster dengan Sistem Informasi Geografis (SIG)

Hasil cluster terbentuk selanjutnya akan dilakukan visualisasi yang berfungsi untuk mempermudah penulis dalam menganalisis kelompok-kelompok DAS berdasarkan parameter karakteristiknya. Dalam penelitian ini visualisasi hasil clustering menggunakan aplikasi sistem informasi geografis (SIG) berbasis Google Map API. Aplikasi SIG akan memplotkan data DAS hasil clustering dengan melibatkan aspek spasialnya (lokasi DAS).

Dalam proses ini data DAS dan hasil clustering DAS yang telah disimpan di database MySQL akan ditampilkan di Google Map menggunakan Pemograman Google Map API. Tahapan dalam proses visualisasi SIG dengan Google Map API adalah sebagai berikut :

1. Pra Prosesing Data

Pada tahap ini dua jenis data yaitu data hasil clustering dan data lokasi DAS beserta atributnya akan digabungkan dengan membuat tabel baru. Tabel yang baru terdiri dari atribut data lokasi DAS dan atribut hasil clustering serta menambahkan atribut koordinat latitude dan longitude. Atribut latitude dan longitude ini yang akan menandai satu lokasi DAS dengan DAS yang lain.

2. Membuat keluaran XML dengan PHP

Google Maps API diintegrasikan dengan database MySQL dan PHP menggunakan file XML sebagai perantara antara database dan Google Map.

3. Menampilkan Peta

Selanjutnya adalah menampilkan XML ke layar melalui script PHP (index.php). Script ini merupakan script utama yang berfungsi untuk menampilkan peta berbasis Google Maps ke dalam *browser*.

[Halaman ini sengaja dikosongkan]

BAB IV

HASIL DAN PEMBAHASAN

4.1 Tahap Pra Prosesing Data

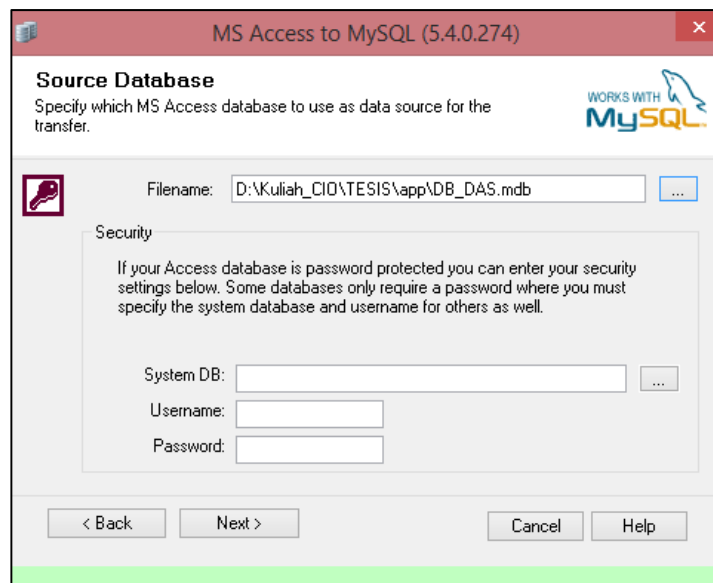
Data awal dalam penelitian ini diperoleh dari Balai Pengelolaan daerah Aliran Sungai (DAS) Tondano Sulawesi Utara. Data ini merupakan data lahan kritis di propinsi sulawesi utara pada tahun 2013 yang terdiri dari 54485 record dan 20 atribut yang disimpan dalam format excel (xlsx).

Tahapan awal dari penelitian ini adalah pra prosesing data. Proses ini meliputi Transformasi data, Pembersihan data dan pemilihan data uji. Dari Proses ini akan menghasilkan data yang siap untuk di lakukan *clustering* di tahap selanjutnya.

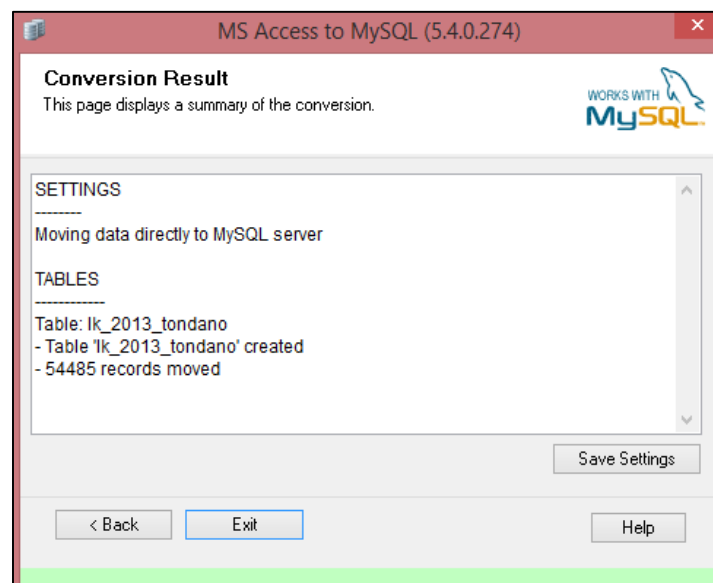
4.1.1 Transformasi Data

Dalam penelitian ini perangkat lunak *Database Management System* (DBMS) yang digunakan untuk menyimpan data DAS dan melakukan Clustering adalah MySQL. Proses transformasi data dilakukan untuk mendapatkan data dalam Format MySQL dari data awal yang disimpan dalam format excel (xlsx).

Langkah pertama yang dilakukan adalah membuat database dengan nama format .mdb melalui MS Access yang berisi data DAS dari format excel (xlsx). Melalui MS Access Data DAS diisi dengan mengimport langsung dari excel (xlsx) ke MS Access. Setelah kita mendapatkan data dalam format .mdb, barulah dilakukan konversi ke bentuk MySQL dengan bantuan perangkat lunak *MS Access to MySQL (v5.4)*. proses konversi data dari format .mdb ke MySQL dan data hasil konversi menggunakan *MS Access to MySQL (v5.4)* ditunjukkan dalam Gambar 4.1 dan 4.2



Gambar 4.1. Konversi data dari format mdb ke MySQL menggunakan *MS Access to MySQL (v5.4)*



Gambar 4.2. Data berhasil dikonversi dari format .mdb ke MySQL menggunakan *MS Access to MySQL (v5.4)*

Dari Gambar 4.2 terlihat bahwa data DAS yang berhasil dikonversi dari MS Access ke MySQL sesuai dengan data awal, yaitu berjumlah 54485 record.

4.1.2 Pembersihan Data dan Pemilihan Data Uji

Setelah data didapat dalam format MySQL, langkah selanjutnya pada tahapan pra prosesing adalah melakukan pembersihan data dan memilih data DAS yang akan dilakukan uji coba. Proses pembersihan data dilakukan dengan tujuan menghilangkan atribut yang tidak dipakai dalam proses clustering dan field data yang kosong atau isi field tidak sesuai dengan kategori. Sedangkan pemilihan data uji dilakukan selain untuk mereduksi jumlah data juga untuk memilih data DAS yang menjadi prioritas untuk dilakukan analisis, dalam penelitian ini akan dipilih DAS dengan luasan diatas 200 Ha. Proses pembersihan data dan pemilihan data uji menggunakan *Query* database ditunjukkan dalam Tabel. 4.1.

Tabel. 4.1 Pembersihan dan Pemilihan data uji dengan *Query*

No.	Keterangan Pembersihan Data Atribut/Field	Script SQL	Jumlah Data Berkurang	Data Uji
1.	Memilih atribut yang dipakai dalam clustering	<i>select KD_DAS, skor_veg, skor_ler, skor_erosi, skor_prod, skor_man, LUAS_HA from tb_das_tondano</i>	0	54485
2.	Menghapus data DAS dengan luasan = 0	<i>select from tb_das_tondano where luas_ha=0</i>	249	54236
3.	Menghapus data DAS dengan kategori tidak sesuai dengan skor parameter karakteristik DAS	<i>Delete tb_das_tondano where skor_veg not in (1,2,3,4,5) and skor_ler not in (1,2,3,4,5) And skor_ero not in (2,3,4,5) And skor_mnj not in (1,3,5) and skor_prod not in (1,2,3,4,5)</i>	141	54095
4.	Menghapus data DAS dengan luasan < 200 Ha	<i>Delete from tb_das_tondano where luas_ha < 200</i>	52958	1137

Dari proses pembersihan data dan pemilihan data uji yang ditunjukkan pada Tabel 4.1, terlihat bahwa data DAS yang siap untuk dilakukan clustering di tahap selanjutnya adalah berjumlah 1137 record.

4.2 Hasil Clustering DAS menggunakan Algoritma *K-Modes Clustering*.

Pada tahapan ini dilakukan clustering DAS menggunakan algoritma k-modes clustering. Aplikasi clustering DAS dengan K-Modes yang dikembangkan dalam penelitian ini menggunakan bahasa pemrograman Visual basic 6.0 yang dikoneksikan dengan database MySQL. Proses dalam algoritma k-modes clustering seperti penentuan cantorid awal, penghitungan jarak data ke centroid dan pengelompokkan data ke dalam cluster dilakukan menggunakan *query* database dan hasilnya akan ditampilkan dalam bentuk aplikasi dengan Visual basic 6.0.

Langkah-langkah Implementasi clustering data DAS menggunakan Visual Basic 6.0 dan MySQL yaitu :


1. Menentukan jumlah cluster yang akan dihasilkan dan inisialisasi centroid (modus) awal untuk setiap cluster. misalkan kita tentukan jumlah cluster yang akan dihasilkan adalah 3 cluster, query database penentuan centroid awal secara random ditunjukkan pada Gambar 4.3.

```
select a.skor_veg as s_veg,a.skor_ler as s_ler,a.skor_erosi
as s_ero,a.skor_prod as s_pro,a.skor_man as s_man
from (select distinct skor_veg,skor_ler,
skor_erosi,skor_prod,skor_man from clus_das_200 group by
skor_veg,skor_ler,skor_erosi,skor_prod,skor_man) a
ORDER BY RAND() LIMIT 0,3
```

Gambar 4.3 Query database penentuan centroid awal secara random

Pada aplikasi yang telah dibuat, centroid (modus) awal yang dihasilkan dari query diatas ditunjukkan pada Gambar 4.4.

Jumlah Cluster : 3

 Proses

Centroid Awal						
	CEN	S_VEG	S_LER	S_ERO	S_PRO	S_MAN
▶	c1	5	4	4	5	5
	c2	3	5	2	3	3
	c3	5	3	4	5	3

Gambar 4.4 Centroid (modus) awal pada aplikasi

2. Menghitung jarak semua data dengan centroid menggunakan formula ketidakmiripan, Query database untuk menghitung jarak data dengan centroid ditunjukkan pada Gambar 4.5.

```
select t.*,
(select a.nama_cen from (select a.kd_das,b.nama_cen,
if (a.skor_veg=b.skor_veg, '0','1') as J_VEG,
if (a.skor_ler=b.skor_ler, '0','1') as J_LER,
if (a.skor_erosi=b.skor_erosi, '0','1') as J_ER0,
if (a.skor_prod=b.skor_prod, '0','1') as J_PRO,
if (a.skor_man=b.skor_man, '0','1') as J_MAN,
(if (a.skor_veg=b.skor_veg, '0','1')+
(if (a.skor_ler=b.skor_ler, '0','1'))+
(if (a.skor_erosi=b.skor_erosi, '0','1'))+
(if (a.skor_prod=b.skor_prod, '0','1'))+
(if (a.skor_man=b.skor_man, '0','1')))) as TOT_J
from clus_das_200 a, centroid b where b.id_cen=1) a
where ((a.kd_das = t.kd_das))
and (a.tot_J = (select min(b.tot_J) from
(select a.kd_das,b.nama_cen,
if (a.skor_veg=b.skor_veg, '0','1') as J_VEG,
if (a.skor_ler=b.skor_ler, '0','1') as J_LER,
if (a.skor_erosi=b.skor_erosi, '0','1') as J_ER0,
if (a.skor_prod=b.skor_prod, '0','1') as J_PRO,
if (a.skor_man=b.skor_man, '0','1') as J_MAN,
(if (a.skor_veg=b.skor_veg, '0','1')+
(if (a.skor_ler=b.skor_ler, '0','1'))+
(if (a.skor_erosi=b.skor_erosi, '0','1'))+
(if (a.skor_prod=b.skor_prod, '0','1'))+
(if (a.skor_man=b.skor_man, '0','1')))) as TOT_J
from clus_das_200 a, centroid b where b.id_cen=1) b
where (b.kd_das = t.kd_das))) limit 1 ) as CLUSTER
from (select a.KD_DAS,a.SKOR_VEG as S_VEG,a.SKOR_LER as
S_LER,a.SKOR_EROSI as S_ER0,a.SKOR_PROD as S_PRO,a.SKOR_MAN as
S_MAN,b.NAMA_CEN AS CEN,if (a.skor_veg=b.skor_veg, '0','1') as
J_VEG,if (a.skor_ler=b.skor_ler, '0','1') as J_LER,
if (a.skor_erosi=b.skor_erosi, '0','1') as J_ER0,
if (a.skor_prod=b.skor_prod, '0','1') as J_PRO,
if (a.skor_man=b.skor_man, '0','1') as J_MAN,
(if (a.skor_veg=b.skor_veg, '0','1')+
(if (a.skor_ler=b.skor_ler, '0','1'))+
(if (a.skor_erosi=b.skor_erosi, '0','1'))+
(if (a.skor_prod=b.skor_prod, '0','1'))+
(if (a.skor_man=b.skor_man, '0','1')))) as TOT_J
from clus_das_200 a, centroid b where b.id_cen=1) t
```

Gambar 4.5 Query untuk menghitung jarak data dengan centroid

Pada aplikasi yang telah dibuat, jarak antara data dengan centroid serta cluster yang diikuti hasil dari query diatas ditunjukkan pada Gambar 4.6

Data DAS Hasil Clustering													
KD_DAS	S_VEG	S_LER	S_ERO	S_PRO	S_MAN	CEN	J_VEG	J_LER	J_ERO	J_PRO	J_MAN	TOT_J	CLUSTER
DAS510130	3	5	3	3	3	c1	1	1	1	1	1	5	c2
DAS510130	3	5	3	3	3	c2	0	0	1	0	0	1	c2
DAS510130	3	5	3	3	3	c3	1	1	1	1	0	4	c2
DAS510353	3	5	2	3	3	c1	1	1	1	1	1	5	c2
DAS510353	3	5	2	3	3	c2	0	0	0	0	0	0	c2
DAS510353	3	5	2	3	3	c3	1	1	1	1	0	4	c2
DAS510471	1	5	3	3	5	c1	1	1	1	1	0	4	c2
DAS510471	1	5	3	3	5	c2	1	0	1	0	1	3	c2
DAS510471	1	5	3	3	5	c3	1	1	1	1	1	5	c2
DAS510487	1	5	5	3	5	c1	1	1	1	1	0	4	c2
DAS510487	1	5	5	3	5	c2	1	0	1	0	1	3	c2
DAS510487	1	5	5	3	5	c3	1	1	1	1	1	5	c2
DAS511260	4	3	3	4	3	c1	1	1	1	1	1	5	c3
DAS511260	4	3	3	4	3	c2	1	1	1	1	0	4	c3
DAS511260	4	3	3	4	3	c3	0	1	1	0	0	2	c3
DAS511266	4	3	3	4	3	c1	1	1	1	1	1	5	c3
DAS511266	4	3	3	4	3	c2	1	1	1	1	0	4	c3
DAS511266	4	3	3	4	3	c3	0	1	1	0	0	2	c3
DAS511365	4	1	4	4	3	c1	1	0	0	1	1	3	c3
DAS511365	4	1	4	4	3	c2	1	1	1	1	0	4	c3

Gambar 4.6 Jarak antara data dengan centorid dan cluster yang diikuti pada aplikasi

- Perbaruhi modulus (sebagai centroid baru) dari setiap cluster berdasarkan data yang bergabung pada setiap clusternya, query database untuk memperoleh centroid (modus) baru ditunjukkan pada Gambar 4.7.

```
select a.cluster,a.s_veg,b.s_ler,c.s_ero,d.s_pro,e.s_man from
(SELECT b.cluster, b.s_veg FROM (SELECT cluster, MAX(f1) as
flmax FROM cen_a GROUP BY cluster) AS a INNER JOIN cen_a AS b
ON a.cluster = b.cluster AND a.flmax = b.f1) a,
(SELECT b.cluster, b.s_ler FROM (SELECT cluster, MAX(f1) as
flmax FROM cen_b GROUP BY cluster) AS a INNER JOIN cen_b AS b
ON a.cluster = b.cluster AND a.flmax = b.f1) b,
(SELECT b.cluster, b.s_ero FROM (SELECT cluster, MAX(f1) as
flmax FROM cen_c GROUP BY cluster) AS a INNER JOIN cen_c AS b
ON a.cluster = b.cluster AND a.flmax = b.f1) c,
(SELECT b.cluster, b.s_pro FROM (SELECT cluster, MAX(f1) as
flmax FROM cen_d GROUP BY cluster) AS a INNER JOIN cen_d AS b
ON a.cluster = b.cluster AND a.flmax = b.f1) d,
(SELECT b.cluster, b.s_man FROM (SELECT cluster, MAX(f1) as
flmax FROM cen_e GROUP BY cluster) AS a INNER JOIN cen_e AS b
ON a.cluster = b.cluster AND a.flmax = b.f1) e
Where a.cluster = b.cluster And a.cluster = c.cluster And
a.cluster = d.cluster And a.cluster = e.cluster group by
a.cluster
```

Gambar 4.7 Query untuk memperoleh centroid (modus) baru

Pada aplikasi yang telah dibuat, centroid (modus) baru yang dihasilkan dari query diatas ditunjukkan pada Gambar 4.8

CEN	S_VEG	S_LER	S_ERD	S_PRO	S_MAN
c1	5	1	4	5	5
c2	3	5	2	3	3
c3	4	1	4	4	3

Gambar 4.8 Centroid (modus) baru

Selanjutnya aplikasi akan mengulang Proses ke-2 dan 3 menggunakan centroid yang baru sampai tidak ada data yang berpindah cluster atau posisi *centroid* tidak berubah. Dalam contoh implementasi ini untuk data yang didapatkan pada masing-masing cluster akhir ditunjukkan pada Gambar 4.9.

Rekap Data		
	CLUSTER	JML_DATA
▶	c1	304
	c2	523
	c3	310

Gambar 4.9 Jumlah data pada masing-masing cluster

4.3 Hasil Validitas Cluster dengan *Davies Bouldin Index*

Setelah proses implementasi clustering menghasilkan cluster dan anggota dari cluster tersebut , poses selanjutnya adalah implementasi penghitungan nilai validitas cluster menggunakan *Davies Bouldin Index* (DBI). Sama seperti proses clustering, penerapan algoritma DBI seperti penghitungan nilai Sum of square within cluster (SSW) , Sum of square between cluster (SSB) dan mencari nilai R_{ij} dilakukan menggunakan *query* database dan hasilnya akan ditampilkan dalam bentuk aplikasi dengan Visual basic 6.0.

Langkah-langkah Implementasi penghitungan nilai DBI menggunakan Visual Basic 6.0 dan MySQL yaitu :

1. Mengelompokan data berdasarkan clusternya dan hitung rata-ratanya untuk menjadi SSW serta menghitung jarak antar centroid untuk menjadi SSB. query database untuk menghitung nilai SSW dan SSB ditunjukkan pada Gambar 4.10.

```

select a.nama_cen,b.ssw,a.tot_jrk as SSB
from (select b.nama_cen,
if (a.skor_veg=b.skor_veg, '0','1') as JRK_VEG,
if (a.skor_ler=b.skor_ler, '0','1') as JRK_LER,
if (a.skor_erosi=b.skor_erosi, '0','1') as JRK_ER0,
if (a.skor_prod=b.skor_prod, '0','1') as JRK_PRO,
if (a.skor_man=b.skor_man, '0','1') as JRK_MAN,
(if (a.skor_veg=b.skor_veg, '0','1')+
(if (a.skor_ler=b.skor_ler, '0','1'))+
(if (a.skor_erosi=b.skor_erosi, '0','1'))+
(if (a.skor_prod=b.skor_prod, '0','1'))+
(if (a.skor_man=b.skor_man, '0','1')))) as TOT_JRK
from centroid_1 a, centroid_1 b) a,
(select CLUSTER,avg(tot_j) as SSW
from x where cen=cluster group by cluster) b
where a.nama_cen=b.cluster

```

Gambar 4.10 Query untuk menghitung nilai SSW dan SSB

Pada aplikasi yang telah dibuat, nilai SSW dan SSB yang dihasilkan dari query diatas ditunjukkan pada Gambar 4.7

Nilai SSW dan SSB		
	CLUS	SSW SSB
	c1	0.78618420
	c1	0.78618425
	c1	0.78618423
	c2	1.59847035
	c2	1.59847030
	c2	1.59847034
►	c3	1.09032253
	c3	1.09032254
	c3	1.09032250

Gambar 4.11 Nilai SSW dan SSB pada aplikasi

- Setelah didapatkan nilai SSW dan SSB, langkah selajutnya adalah menghitung nilai R_{ij} . Sebelum melakukan query untuk menghitung nilai R_{ij} , terlebih dahulu dibuat tabel ssw_ssb yang berisi hasil dari query mencari nilai SSW dan SSB sebelumnya. query database untuk menghitung nilai R_{ij} ditunjukkan pada Gambar 4.12.

```

select a.nama_cen as CLUST,a.SSW,a.cluster as
CEN,a.ssw1 as SSW_CLUS,b.tot_jrk as SSB,
sum((a.ssw+a.ssw1)/b.tot_jrk) as NILAI_R
from r a,ssw_ssb b where a.id_r=b.id_s
group by a.nama_cen,a.ssw,a.cluster,a.ssw1,b.tot_jrk

```

Gambar 4.12 Query untuk menghitung nilai R_{ij}

Pada aplikasi yang telah dibuat, rekapitulasi nilai SSW, SSB dan R_{ij} yang dihasilkan dari query diatas ditunjukkan pada Gambar 4.13

Nilai SSW dan SSB Davies-Bouldin Index						
	CLUS	SSW	CEN	SSW_CL	SSB	NILAI_R
	c1	0.7861842	c1	0.7861840		
	c1	0.7861842	c2	1.5984705		0.4769309
	c1	0.7861842	c3	1.0903223		0.6255022
	c2	1.5984703	c1	0.7861845		0.4769309
	c2	1.5984703	c2	1.5984700		
	c2	1.5984703	c3	1.0903224		0.6721982
►	c3	1.0903225	c1	0.7861843		0.6255022
	c3	1.0903225	c2	1.5984704		0.6721982
	c3	1.0903225	c3	1.0903220		

Gambar 4.13 Rekap nilai SSW, SSB dan R_{ij} pada aplikasi

- Langkah terakhir adalah menghitung nilai DBI, Dimana $R_{ij \text{ max}}$ pada masing cluster akan dirata-rata yang merupakan nilai DBI akhir. query database untuk menghitung nilai $R_{ij \text{ max}}$ ditunjukkan pada Gambar 4.14.

```
select a.nama_cen as CLUSTER,
max((a.ssw+a.ssw1)/b.tot_jrk)
as R_MAX from r a,ssw_ssb b where
a.id_r=b.id_s group by a.nama_cen
```

Gambar 4.14 Query untuk menghitung nilai $R_{ij \text{ max}}$

Pada aplikasi yang telah dibuat, nilai $R_{ij \text{ max}}$ yang dihasilkan dari query diatas ditunjukkan pada Gambar 4.9.

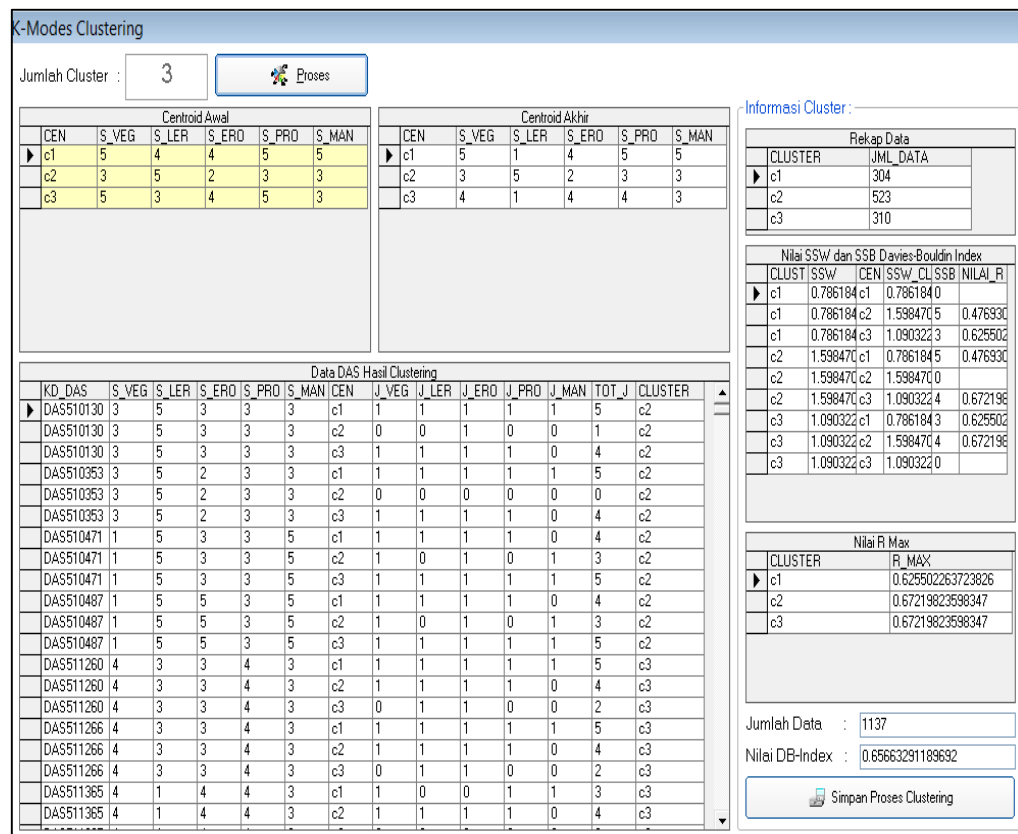
Nilai R Max	
CLUSTER	R_MAX
c1	0.625502263723826
c2	0.67219823598347
► c3	0.67219823598347

Gambar 4.15 Nilai $R_{ij \text{ max}}$ pada aplikasi

Sehingga dari implementasi Terlihat bahwa nilai DBI yang didapatkan adalah rata-rata dari nilai $R_{ij \text{ max}}$ untuk masing-masing cluster atau 0.657.

4.4 Hasil Analisis dan Optimasi Clustering

Dari implementasi percobaan menggunakan 3 cluster di atas, proses clustering DAS secara keseluruhan dapat dilihat pada Gambar 4.16



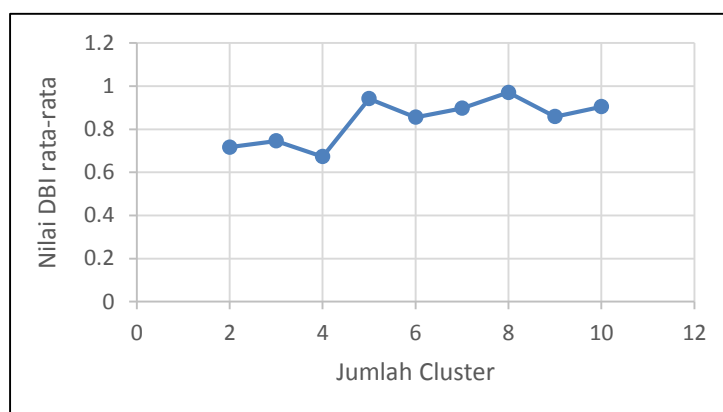
Gambar 4.16 Proses clustering DAS secara keseluruhan

Optimasi cluster dilakukan untuk mengetahui berapa jumlah cluster yang menghasilkan nilai DBI terendah, artinya jumlah cluster dengan tingkat kemiripan data yang paling mirip. Dari hasil percobaan menggunakan 2 sampai 10 cluster dengan nilai *centroid* awal secara random sebanyak 10 kali, Rekapitulasi nilai DBI yang didapatkan ditunjukkan dalam Tabel 4.2.

Tabel. 4.2 Rekapitulasi nilai DBI

Jml Cluster	Jml Testing	DBI Min	DBI Max	DBI Rata-rata
2	10	0.707836	0.723965	0.716214
3	10	0.604369	1.375525	0.745669
4	10	0.596779	0.876039	0.672778
5	10	0.617916	1.577397	0.941772
6	10	0.657734	1.227203	0.855413
7	10	0.723873	1.048321	0.896865
8	10	0.693813	1.374299	0.97061
9	10	0.719653	1.217179	0.858375
10	10	0.69299	1.327828	0.904521
Nilai DBI rata-rata				0.840246

Dari tabel rekapitulasi nilai DBI didapatkan grafik nilai DBI rata-rata untuk 2 sampai 10 cluster pada Gambar 4.17.



Gambar 4.17 Grafik nilai DBI rata-rata terhadap jumlah cluster

Pada Tabel 4.2 dapat dilihat bahwa nilai DBI rata-rata untuk semua percobaan adalah 0.840246, sedangkan jumlah cluster dengan rata-rata nilai DBI terendah adalah 4 cluster yaitu dengan nilai DBI rata-rata 0.672778 atau lebih kecil 19,93% dari nilai DBI rata-rata untuk semua percobaan. Dari percobaan dengan 2 sampai 10 cluster didapatkan jumlah cluster

optimal yang akan dijadikan jumlah cluster dalam *clustering* data DAS adalah 4 cluster.

Hasil *clustering* menggunakan 4 cluster dengan nilai *centroid* awal secara random sebanyak 10 kali ditunjukkan pada Tabel 4.3.

Tabel 4.3 Hasil clustering dengan 4 cluster

Percobaan Ke-	Id_Coba	Jml Cluster	Nilai DBI
1	199	4	0.636312417
2	207	4	0.677207338
3	208	4	0.816820806
4	213	4	0.632737921
5	214	4	0.607686581
6	269	4	0.609031623
7	270	4	0.876038618
8	271	4	0.636303181
9	272	4	0.596779152
10	273	4	0.638863456

Dari hasil clustering dengan 4 cluster pada Tabel 4.3, nilai DBI terendah didapat pada percobaan ke-9 dengan nilai DBI 0.596779152. sehingga percobaan ke-9 dengan 4 cluster akan dijadikan clustering optimal dalam pengelompokan DAS berdasarkan parameter karakteristik DAS. Detail hasil clustering pada percobaan ke-9 dengan 4 cluster pada aplikasi ditunjukkan pada Gambar 4.18.

Detail Hasil Clustering DAS

Id Percobaan : 272

Jml Cluster : 4

Nilai DB-Index : 0.596779152238595

Peta Hasil Clustering

Jumlah Data		Centroid Awal						Centroid Akhir					
CLUSTER	JML DATA	CEN	S_VEG	S_LER	S_ERD	S_PROCS	MAN	CEN	S_VEG	S_LER	S_ERD	S_PROCS	MAN
c1	348	c1	5	5	5	5	5	c1	5	1	4	5	5
c2	269	c2	2	2	3	2	3	c2	4	2	4	4	3
c3	332	c3	3	5	2	3	3	c3	3	2	2	3	3
c4	188	c4	1	2	3	3	3	c4	1	5	5	3	3

Data DAS Hasil Clustering						
KD_DAS	DAS	KAB_KOTA	KECAMATAN	DESA	LUAS_HA	CLUSTER
DAS510130	Abuang	MINAHASA TENGGARA	BELANG	MOLOMPAR UTARA	239.5474	c3
DAS510353	Akas	KEPULAUAN TALAUD	DAMAU	DAMAU BAWONE	472.7836	c3
DAS510471	Alung Banua	MANADO	BUNAKEN	ALUNG BANUA	255.3915	c4
DAS510487	Alung Banua	MANADO	BUNAKEN	BUNAKEN	282.1592	c4
DAS511260	Atinggola	BOLAANG MONGONDOW UTARA	PINGGALUMAN	TUNTUNG	206.7353	c2
DAS511266	Atinggola	BOLAANG MONGONDOW UTARA	PINGGALUMAN	TUNTUNG TIMUR	201.294	c2
DAS511365	Atinggola	BOLAANG MONGONDOW UTARA	PINGGALUMAN	TUNTUNG TIMUR	205.3994	c2
DAS511403	Awit	KEPULAUAN TALAUD	BEO UTARA	AWIT SELATAN	454.1657	c2
DAS511404	Awit	KEPULAUAN TALAUD	BEO UTARA	RAE	375.5362	c2
DAS511432	Awit 2	KEPULAUAN TALAUD	BEO UTARA	RAE	219.7818	c2
DAS511491	Ayong	BOLAANG MONGONDOW	SANGTOMBOLANG	BOLANGAT	889.4589	c2
DAS511559	Ayong	BOLAANG MONGONDOW	SANGTOMBOLANG	BABO	255.787	c2
DAS511563	Ayong	BOLAANG MONGONDOW	SANGTOMBOLANG	BOLANGAT	446.2193	c2
DAS511566	Ayong	BOLAANG MONGONDOW	SANGTOMBOLANG	CEMPAKA	227.3713	c2
DAS511576	Ayong	BOLAANG MONGONDOW	LOLAK	BUNTALO TIMUR	330.8437	c1
DAS511580	Awana	BOLAANG MONGONDOW	SANGTOMBOLANG	BABO	228.3759	c1

Gambar 4.18 Detail hasil clustering akhir DAS

Terlihat dari detail hasil akhir clustering dalam pengelompokan DAS berdasarkan parameter karakteristik DAS pada percobaan ke-9 dengan 4 cluster, cluster dengan anggota paling banyak adalah cluster 1 dengan 348 DAS, Sedangkan cluster dengan anggota paling sedikit adalah cluster 4 dengan 188 DAS. Profiling karakteristik DAS yang terbentuk pada masing-masing cluster/segmen dari Proses clustering pada percobaan ke-9 dengan 4 cluster ditunjukkan pada Tabel 4.4.

Tabel 4.4. Karakteristik DAS pada setiap cluster/segmen

Cluster	Jml Data	%	Karakteristik DAS
1	348	30,6 %	Kelompok DAS dengan karakteristik penutupan lahan yang sangat baik dan tingkat kelerengan yang sangat curam. Tingkat erosi untuk DAS di cluster ini sedang dengan tingkat produktivitas lahan sangat tinggi dan manajemen yang baik.
2	269	23,7 %	Kelompok DAS dengan penutupan lahan yang baik dan tingkat kelerengan yang curam. Tingkat erosi berada pada tingkat yang sedang dengan tingkat produktivitas tinggi serta manajemen pada tingkat sedang.
3	332	29,2 %	Kelompok DAS dengan karakteristik tingkat penutupan lahan sedang dan tingkat kelerengan yang curam. Tingkat erosi pada kelompok ini berada pada tingkat yang berat dengan produktivitas lahan sedang serta manajemen pada tingkat sedang.
4	188	16,5 %	Kelompok DAS dengan tingkat penutupan lahan yang sangat buruk namun memiliki tingkat kelerengan yang datar. Tingkat erosi DAS di kelompok ini ringan dengan tingkat produktivitas lahan sedang serta manajemen pada tingkat sedang.

Sementara sebaran DAS yang didapatkan dari Proses clustering pada percobaan ke-9 dengan 4 cluster untuk Kab/Kota pada masing- masing cluster ditunjukkan pada Tabel 4.5.

Tabel 4.5. Sebaran DAS untuk Kab/Kota pada masing-masing cluster/segmen

Kab/Kota	Jumlah data DAS			
	Cluster 1	Cluster 2	Cluster 3	Cluster 4
BITUNG	1 (0.3%)	8 (3%)	9 (2.7%)	1 (0.5%)
BOLMONG	125 (35.9%)	86 (32%)	47 (14.2%)	55 (29.3%)
BOLMONG SELATAN	114 (32.8%)	35 (13%)	17 (5.1%)	7 (3.7%)
BOLMONG TIMUR	35 (10.1%)	21 (7.8%)	17 (5.1%)	10 (5.3%)
BOLMONG UTARA	44 (12.6%)	75 (27.9%)	14 (4.2%)	7 (3.7%)
KEPULAUAN SANGIHE	1 (0.3%)	1 (0.4%)	18 (5.4%)	3 (1.6%)
KEPULAUAN TALAUD	2 (0.6%)	20 (7.4%)	16 (4.8%)	2 (1.1%)
KOTAMOBAGU	0 (0%)	0 (0%)	0 (0%)	1 (0.5%)
MANADO	0 (0%)	0 (0%)	5 (1.5%)	7 (3.7%)
MINAHASA	1 (0.3%)	4 (1.5%)	35 (10.5%)	25 (13.3%)
MINAHASA SELATAN	17 (4.9%)	6 (2.2%)	82 (24.7%)	17 (9.0%)
MINAHASA TENGGARA	4 (1.1%)	2 (0.7%)	43 (13%)	15 (8.0%)
MINAHASA UTARA	4 (1.1%)	11 (4.1%)	24 (7.2%)	35 (18.6%)
TOMOHON	0 (0%)	0 (0%)	5 (1.5%)	3 (1.6%)
Total Data	348	269	332	188

Uji akurasi data hasil clustering dengan sistem perhitungan manual yang dipakai untuk menentukan tingkat kekritisan lahan selama ini dilakukan dengan membandingkan Empat cluster yang terbentuk dengan 5 kategori tingkat kekritisan lahan yang menjadi acuan dalam menentukan prioritas rehabilitasi DAS. Perbandingan jumlah DAS hasil clustering dengan penghitungan manual berdasarkan tingkat kekritisan lahan ditunjukkan pada Tabel 4.6.

Tabel 4.6 Perbandingan jumlah DAS hasil clustering dengan penghitungan manual tingkat kekritisan lahan.

Hasil Cluster/ Segmen dengan K- Modes	Kategori Kritis dengan Perhitungan Manual					
	Sangat Kritis	Kritis	Agak Kritis	Potensial Kritis	Tidak Kritis	<i>Jumlah</i> DAS
c1	0	5	51	278	14	348
c2	0	6	158	105	0	269
c3	10	145	174	3	0	332
c4	1	4	38	144	1	188
<i>Jumlah</i> DAS	11	160	421	530	15	1137

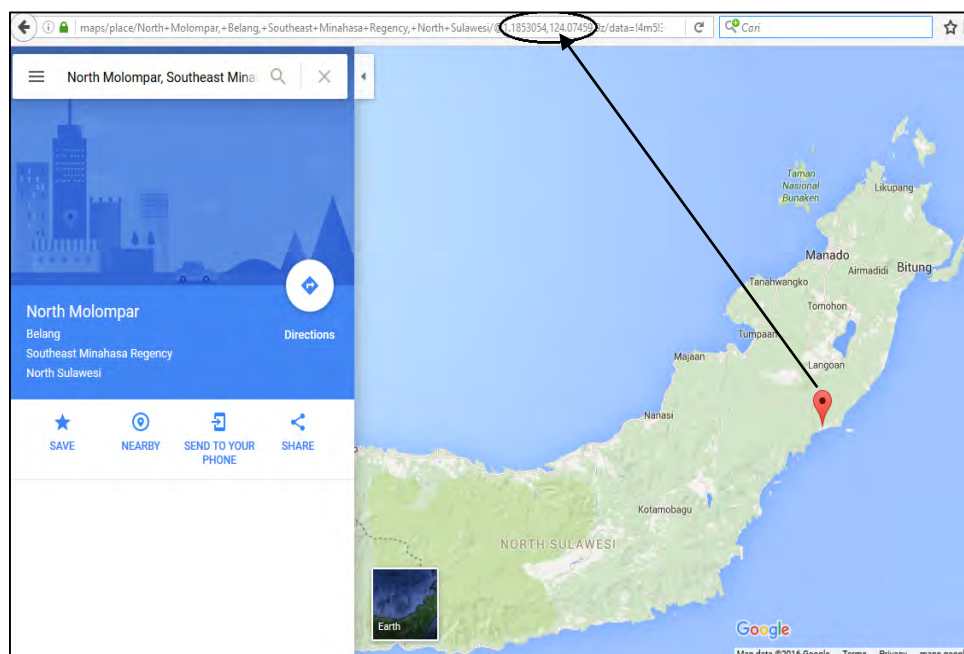
Pada analisa profiling DAS untuk masing-masing cluster/segmen sebelumnya, cluster 3 adalah cluster dengan tingkat kekritisan lahan tertinggi dibandingkan dengan cluster lain. Apabila dibandingkan dengan penghitungan manual untuk kategori DAS dengan prioritas rehabilitasi pada Tabel 4.5, DAS dengan kekritisan lahan tertinggi atau pada kategori sangat kritis dan kritis dengan jumlah 171 DAS didominasi DAS di cluster 3, yaitu 155 DAS atau 90,64 %. Ini berarti hasil dari proses clustering tidak jauh berbeda apabila dibandingkan dengan penghitungan tingkat kekritisan DAS secara manual .

4.5 Hasil Visualisasi Clustering DAS dengan Google Map API

Tahapan akhir dalam penelitian ini adalah visualisasi hasil clustering dengan Sistem Informasi Geografis (SIG). Aplikasi SIG akan memplotkan data DAS hasil clustering dengan melibatkan aspek spasialnya (lokasi DAS). Dalam proses ini hasil clustering DAS yang telah disimpan di database MySQL akan ditampilkan di Google Map menggunakan Pemograman Google Map API.

Tahapan dalam proses visualisasi SIG dengan Google Map API yaitu :

1. Data hasil clustering dan data lokasi DAS beserta atributnya akan digabungkan dengan membuat tabel baru dengan nama plot_das. Tabel plot_das terdiri dari atribut data lokasi DAS dan atribut hasil clustering serta menambahkan atribut koordinat latitude dan longitude. Koordinat latitude dan longitude didapatkan dari Google Map dengan melakukan pencarian terhadap atribut lokasi DAS seperti Desa, Kecamatan dan Kabupaten. Proses mendapatkan koordinat latitude dan longitude dari Google Map ditunjukkan pada Gambar 4.19.



Gambar 4.19 Koordinat Latitude dan Longitude Desa Malompar
Utara, Kec. Belang, Kab. Minahasa Tenggara

Proses mendapatkan koordinat latitude dan longitude pada Gambar 4.13 dilakukan untuk semua data DAS. Sehingga didapatkan data DAS dan koordinat lokasi pada tabel plot_das secara lengkap pada Gambar 4.20.

id	nama	lokasi	lat	lng	clus
1120	DAS563007	TOULIANG OKI,ERIS,MINAHASA	1.260951	124.948563	c3
1121	DAS563039	SULUAN,TOMBULU,MINAHASA	1.338040	124.901237	c3
1122	DAS563043	MASARANG,TONDANO BARAT,MINAHASA	1.318101	124.879166	c3
1123	DAS563150	SULUAN,TOMBULU,MINAHASA	1.353341	124.882637	c2
1124	DAS563203	AIRMADIDI ATAS,AIRMADIDI,MINAHASA UTARA	1.451326	124.974022	c2
1125	DAS563268	TANGGARI,AIRMADIDI,MINAHASA UTARA	1.364484	124.942841	c3
1126	DAS563614	TULE,MELONGUANE TIMUR,KEPULAUAN TALAUD	4.027248	126.749367	c3
1127	DAS563713	KAYUBESI,KOMBI,MINAHASA	1.211294	124.970482	c3
1128	DAS563865	TUNTUNG TIMUR,PINOGALUMAN,BOLAANG MONGONDOW UTARA	0.895723	123.205765	c3
1129	DAS564149	ERIS,ERIS,MINAHASA	1.222691	124.940453	c3
1130	DAS564203	BERINGIN,BELANG,MINAHASA TENGGARA	0.965493	124.763512	c4
1131	DAS564248	MOREAH SATU,RATATOTOK,MINAHASA TENGGARA	0.964989	124.727936	c3
1132	DAS564263	BERINGIN,BELANG,MINAHASA TENGGARA	0.966911	124.765373	c3
1133	DAS564269	MOREAH SATU,RATATOTOK,MINAHASA TENGGARA	0.955992	124.729080	c3
1134	DAS564290	MOREAH SATU,RATATOTOK,MINAHASA TENGGARA	0.948073	124.731956	c3
1135	DAS564374	BATU,LIKUPANG SELATAN,MINAHASA UTARA	1.673197	125.026459	c3
1136	DAS564387	WEROT,LIKUPANG SELATAN,MINAHASA UTARA	1.657782	124.997520	c4
1137	DAS564503	SAKTI,POSIGADAN,BOLAANG MONGONDOW SELATAN	0.317792	123.658562	c1

Gambar 4.20 Data pada Tabel Plot_das

2. Menggabungkan Google Maps API dengan database MySQL dan PHP menggunakan file XML. Pertama kita harus membuat koneksi database dengan PHP dan diberi nama koneksi.php (script koneksi.php secara lengkap ada pada lampiran 1). Kemudian menggunakan XML sebagai perantara antara MySQL dengan Google Map dan diberi nama generatexml.php (script generatexml.php secara lengkap ada pada lampiran 2). Jika script generatexml.php dijalankan maka *browser* akan menampilkan data XML seperti pada Gambar 4.21. isi dari data XML tersebut tidak lain adalah isi database MySQL yang kita buat pada langkah sebelumnya.

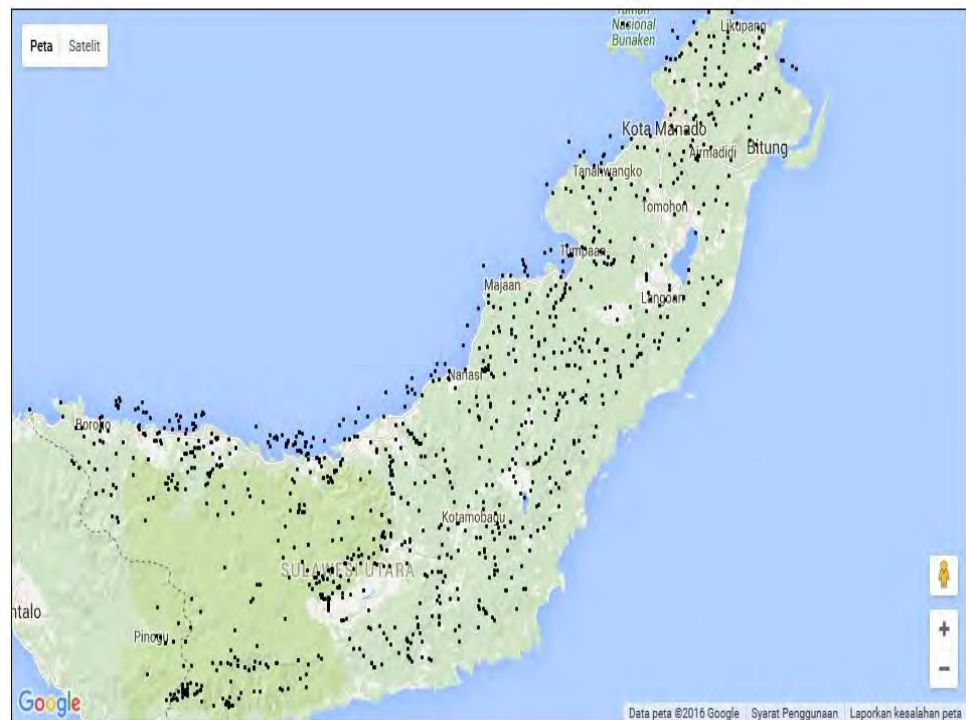
```

localhost/peta/generatexml.php
<marker nama="DASS10487" lokasi="BUNAKEN,BUNAKEN,MANADO" lat="1.622543" lng="124.753326" clus="c4"/>
<marker nama="DASS11260" lokasi="TUNTUNG,PINO GALUMAN,BOLAANG MONGONDOW UTARA" lat="0.868780" lng="123.212425" clus="c2"/>
<marker nama="DASS11266" lokasi="TUNTUNG TIMUR,PINO GALUMAN,BOLAANG MONGONDOW UTARA" lat="0.846729" lng="123.264633" clus="c2"/>
<marker nama="DASS11365" lokasi="TUNTUNG TIMUR,PINO GALUMAN,BOLAANG MONGONDOW UTARA" lat="0.828191" lng="123.268410" clus="c2"/>
<marker nama="DASS11403" lokasi="AWIT SELATAN,BEO UTARA,KEPULAUAN TALAUD" lat="4.355687" lng="126.705490" clus="c2"/>
<marker nama="DASS11404" lokasi="RAE,BEO UTARA,KEPULAUAN TALAUD" lat="4.333741" lng="126.714119" clus="c2"/>
<marker nama="DASS11432" lokasi="RAE,BEO UTARA,KEPULAUAN TALAUD" lat="4.348719" lng="126.765450" clus="c2"/>
<marker nama="DASS11491" lokasi="BOLANGAT,SANGTOMBOLANG,BOLAANG MONGONDOW" lat="0.816514" lng="123.845940" clus="c2"/>
<marker nama="DASS11559" lokasi="BABO,SANGTOMBOLANG,BOLAANG MONGONDOW" lat="0.703429" lng="123.657028" clus="c2"/>
<marker nama="DASS11563" lokasi="BOLANGAT,SANGTOMBOLANG,BOLAANG MONGONDOW" lat="0.802876" lng="123.850883" clus="c2"/>
<marker nama="DASS11566" lokasi="CEMPAKA,SANGTOMBOLANG,BOLAANG MONGONDOW" lat="0.682247" lng="123.675995" clus="c2"/>
<marker nama="DASS11576" lokasi="BUNTALO TIMUR,LOLAK,BOLAANG MONGONDOW" lat="0.740772" lng="123.844116" clus="c1"/>
<marker nama="DASS11580" lokasi="BABO,SANGTOMBOLANG,BOLAANG MONGONDOW" lat="0.819118" lng="123.875381" clus="c1"/>
<marker nama="DASS11606" lokasi="BUNTALO TIMUR,LOLAK,BOLAANG MONGONDOW" lat="0.822146" lng="123.925720" clus="c2"/>
<marker nama="DASS11615" lokasi="BABO,SANGTOMBOLANG,BOLAANG MONGONDOW" lat="0.810193" lng="123.874695" clus="c2"/>
<marker nama="DASS11619" lokasi="BOLANGAT,SANGTOMBOLANG,BOLAANG MONGONDOW" lat="0.814548" lng="123.847450" clus="c2"/>
<marker nama="DASS11621" lokasi="CEMPAKA,SANGTOMBOLANG,BOLAANG MONGONDOW" lat="0.683277" lng="123.853836" clus="c2"/>
<marker nama="DASS11622" lokasi="CEMPAKA,SANGTOMBOLANG,BOLAANG MONGONDOW" lat="0.683706" lng="123.854210" clus="c2"/>
<marker nama="DASS11625" lokasi="CEMPAKA,SANGTOMBOLANG,BOLAANG MONGONDOW" lat="0.697009" lng="123.857269" clus="c2"/>
<marker nama="DASS11645" lokasi="BUNTALO TIMUR,LOLAK,BOLAANG MONGONDOW" lat="0.741974" lng="123.947121" clus="c1"/>
<marker nama="DASS11647" lokasi="BUNTALO TIMUR,LOLAK,BOLAANG MONGONDOW" lat="0.740600" lng="123.937843" clus="c1"/>
<marker nama="DASS11648" lokasi="BUNTALO TIMUR,LOLAK,BOLAANG MONGONDOW" lat="0.744033" lng="123.945396" clus="c1"/>

```

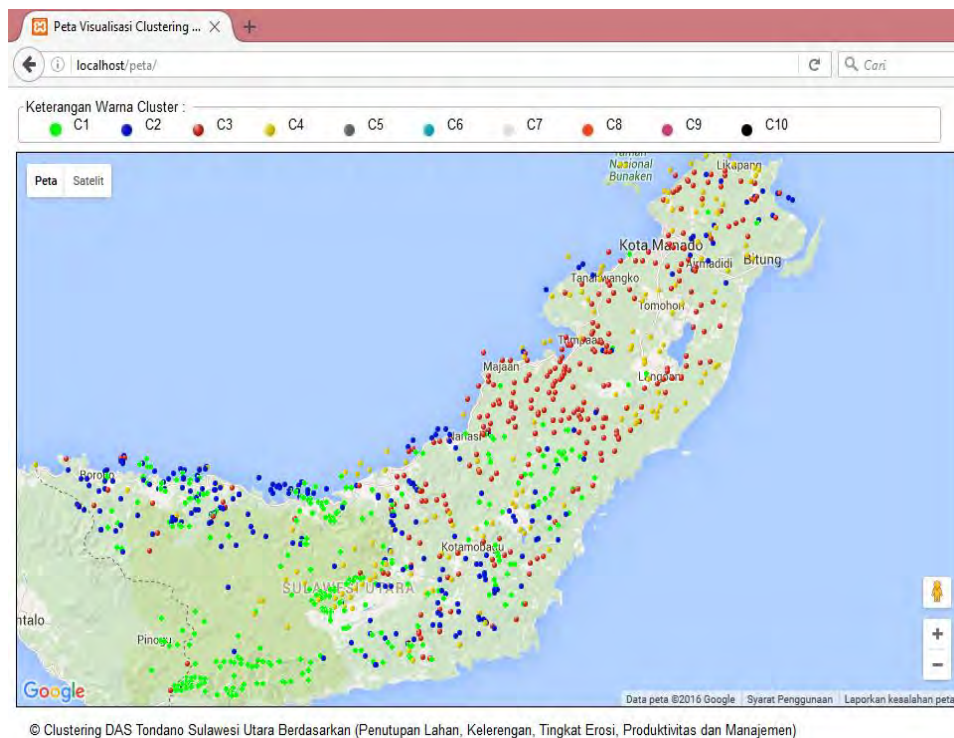
Gambar 4.21 Tampilan XML yang berisi data lokasi DAS

- Menampilkan XML ke layar melalui script index.php (script index.php secara lengkap ada pada lampiran 3). Script ini merupakan script utama yang berfungsi untuk menampilkan peta berbasis Google Maps ke dalam *browser*. Hasil marking semua lokasi DAS di table plot_das yang ditampilkan dalam peta dapat dilihat pada Gambar 4.22.



Gambar 4.22 Hasil marking semua lokasi DAS yang ditampilkan dalam Google Map

Hasil akhir peta yang dihasilkan menampilkan plot hasil clustering dengan data lokasinya. Masing-masing cluster diplotkan dengan warna yang berbeda beda sehingga keterangan setiap cluster dapat dilihat berdasarkan warnanya. Untuk cluster 1 diplotkan dengan warna hijau, cluster 2 dengan warna biru, cluster 3 dengan warna merah dan cluster 4 dengan warna kuning. Visualisasi clustering DAS optimal pada percobaan ke-9 dengan 4 cluster dapat dilihat pada Gambar 4.23.



Gambar 4.23 Visualisasi clustering DAS Tondano berdasarkan parameter karakteristik.

LAMPIRAN 1

Script koneksi database dengan PHP (koneksi.php)

```
<?php
//koneksi ke server
$server="localhost";
$username="root";
$password="";
$konek=mysql_connect($server,$username,$password);
//cek koneksi
if(!$konek){
echo "Koneksi Gagal";
}
//memilih database
$db=mysql_select_db("dbdas");
//cek database
if(!$db){
echo "database Gagal";
}
?>
```

LAMPIRAN 2

Script XML sebagai perantara antara MySQL dengan Google Map (generatexml.php)

```
<?php
require("koneksi.php");
function parseToXML($htmlStr)
{
$xmlStr=str_replace('<','&lt;',$htmlStr);
$xmlStr=str_replace('>','&gt;',$xmlStr);
$xmlStr=str_replace('"','&quot;',$xmlStr);
$xmlStr=str_replace("'",'&#39;',$xmlStr);
$xmlStr=str_replace("&","&amp;",$xmlStr);
return $xmlStr;
}
$connection=mysql_connect (localhost, $username, $password);
if (!$connection) {
die('Not connected : ' . mysql_error());
}
// Mengeset Database MySQL aktif
$db_selected = mysql_select_db("dbdas");
if (!$db_selected) {
die ('Can\'t use db : ' . mysql_error());
}
// Memilih semua baris pada tabel
$query = "SELECT * FROM plot_das WHERE 1";
$result = mysql_query($query);
if (!$result) {
die('Invalid query: ' . mysql_error());
}
header("Content-type: text/xml");
// Mulai XML file, menampilkan parent node
echo '<markers>';
// Iterasi melalui baris, node pencetakan XML untuk setiap baris
while ($row = @mysql_fetch_assoc($result)){
// Menambahkan ke XML document node
echo '<marker ' ;
echo 'nama="' . parseToXML($row['nama']) . '" ' ;
echo 'lokasi="' . parseToXML($row['lokasi']) . '" ' ;
echo 'lat="' . $row['lat'] . '" ' ;
echo 'lng="' . $row['lng'] . '" ' ;
echo 'clus="' . $row['clus'] . '" ' ;
echo '>';
}
echo '</markers>'; ?>
```

LAMPIRAN 3

Script Menampilkan XML ke browser (script index.php)

```
<?php
include "header.php";
?>
<head> <meta http-equiv="content-type" content="text/html; charset=utf-8"/>
<title> Clustering DAS Tondano </title> <script
src="http://maps.google.com/maps?file=api&v=2&sensor=true&key=ABQIAAA8tt4eKTuBZMVnLJfP2BZrBT2yXp_ZAY8_ufC3CFXhHIE1Nv
wkxS4Rz1LFzG0odNPtk8VLkdrQF5grA" type="text/javascript"></script>
<script type="text/javascript"> //
var iconURLPrefix = 'images/';
var icon1 = new GIcon();
icon1.image = iconURLPrefix + 'marker1.png';
icon1.shadow =
'http://labs.google.com/ridefinder/images/mm_20_shadow.png';
icon1.iconSize = new GSize(4, 4);
icon1.shadowSize = new GSize(22, 20);
icon1.iconAnchor = new GPoint(6, 20);
icon1.infoWindowAnchor = new GPoint(5, 1);
var icon2 = new GIcon();
icon2.image = iconURLPrefix + 'marker2.png';
icon2.shadow =
'http://labs.google.com/ridefinder/images/mm_20_shadow.png';
icon2.iconSize = new GSize(4, 4);
icon2.shadowSize = new GSize(22, 20);
icon2.iconAnchor = new GPoint(6, 20);
icon2.infoWindowAnchor = new GPoint(5, 1);
var icon3 = new GIcon();
icon3.image = iconURLPrefix + 'marker3.png';
icon3.shadow =
'http://labs.google.com/ridefinder/images/mm_20_shadow.png';
icon3.iconSize = new GSize(4, 4);
icon3.shadowSize = new GSize(22, 20);
icon3.iconAnchor = new GPoint(6, 20);
icon3.infoWindowAnchor = new GPoint(5, 1);
var icon4 = new GIcon();
icon4.image = iconURLPrefix + 'marker4.png';
icon4.shadow =
'http://labs.google.com/ridefinder/images/mm_20_shadow.png';
icon4.iconSize = new GSize(4, 4);
icon4.shadowSize = new GSize(22, 20);
icon4.iconAnchor = new GPoint(6, 20);
icon4.infoWindowAnchor = new GPoint(5, 1);
var icon5 = new GIcon();</pre></div>
```

```

    icon5.image = iconURLPrefix + 'marker5.png';
    icon5.shadow =
'http://labs.google.com/ridefinder/images/mm_20_shadow.png';
    icon5.iconSize = new GSize(4, 4);
    icon5.shadowSize = new GSize(22, 20);
    icon5.iconAnchor = new GPoint(6, 20);
    icon5.infoWindowAnchor = new GPoint(5, 1);
    var icon6 = new GIcon();
    icon6.image = iconURLPrefix + 'marker6.png';
    icon6.shadow =
'http://labs.google.com/ridefinder/images/mm_20_shadow.png';
    icon6.iconSize = new GSize(4, 4);
    icon6.shadowSize = new GSize(22, 20);
    icon6.iconAnchor = new GPoint(6, 20);
    icon6.infoWindowAnchor = new GPoint(5, 1);
    var icon7 = new GIcon();
    icon7.image = iconURLPrefix + 'marker7.png';
    icon7.shadow =
'http://labs.google.com/ridefinder/images/mm_20_shadow.png';
    icon7.iconSize = new GSize(4, 4);
    icon7.shadowSize = new GSize(22, 20);
    icon7.iconAnchor = new GPoint(6, 20);
    icon7.infoWindowAnchor = new GPoint(5, 1);
    var icon8 = new GIcon();
    icon8.image = iconURLPrefix + 'marker8.png';
    icon8.shadow =
'http://labs.google.com/ridefinder/images/mm_20_shadow.png';
    icon8.iconSize = new GSize(4, 4);
    icon8.shadowSize = new GSize(22, 20);
    icon8.iconAnchor = new GPoint(6, 20);
    icon8.infoWindowAnchor = new GPoint(5, 1);
    var icon9 = new GIcon();
    icon9.image = iconURLPrefix + 'marker9.png';
    icon9.shadow =
'http://labs.google.com/ridefinder/images/mm_20_shadow.png';
    icon9.iconSize = new GSize(4, 4);
    icon9.shadowSize = new GSize(22, 20);
    icon9.iconAnchor = new GPoint(6, 20);
    icon9.infoWindowAnchor = new GPoint(5, 1);
    var icon10 = new GIcon();
    icon10.image = iconURLPrefix + 'marker10.png';
    icon10.shadow =
'http://labs.google.com/ridefinder/images/mm_20_shadow.png';
    icon10.iconSize = new GSize(4, 4);
    icon10.shadowSize = new GSize(22, 20);
    icon10.iconAnchor = new GPoint(6, 20);
    icon10.infoWindowAnchor = new GPoint(5, 1);

```



```

var customIcons = [];
customIcons["c1"] = icon1;
customIcons["c2"] = icon2;
customIcons["c3"] = icon3;
customIcons["c4"] = icon4;
customIcons["c5"] = icon5;
customIcons["c6"] = icon6;
customIcons["c7"] = icon7;
customIcons["c8"] = icon8;
customIcons["c9"] = icon9;
customIcons["c10"] = icon10;
function load() {
if (GBrowserIsCompatible()) {
var map = new GMap2(document.getElementById("map"));
map.addControl(new GSmallMapControl());
map.addControl(new GMapTypeControl());
map.setCenter(new GLatLng(1.0157359, 124.3650567), 9);
GDownloadUrl("generatexml.php", function(data) {
var xml = GXml.parse(data);
var markers = xml.documentElement.getElementsByTagName("marker");
for (var i = 0; i < markers.length; i++) {
var nama = markers[i].getAttribute("nama");
var lokasi = markers[i].getAttribute("lokasi");
var clus = markers[i].getAttribute("clus");
var point = new GLatLng(parseFloat(markers[i].getAttribute("lat")),
parseFloat(markers[i].getAttribute("lng")));
var marker = createMarker(point, nama, lokasi, clus);
map.addOverlay(marker);
} }); }
}
function createMarker(point, nama, lokasi, clus) {
var marker = new GMarker(point, customIcons[clus]);
var html = "<b>" + nama + "</b> <br/>" + lokasi;
GEvent.addListener(marker, 'click', function() {
marker.openInfoWindowHtml(html);
});
return marker; } //]]>
</script>
</head>
<body onload="load()" onunload="GUnload()">
<div id="map" style="width: 950px; height:500px; border: 1px solid black">
<div align="center">
<p>&nbsp;</p>
<p align="left"></p>
</div> </div>
<?php
include "footer.php"; ?> </body> </html>

```

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dengan teknik data mining, Jumlah DAS yang menjadi prioritas rehabilitasi lebih sedikit daripada perhitungan manual. Jika perhitungan manual ada 171 DAS yang harus direhabilitasi maka dengan teknik data mining hanya 155 DAS yang harus direhabilitasi atau 9,36 % lebih sedikit dari perhitungan manual. Dari penelitian ini diharapkan bisa membantu pemerintah dalam hal ini Kementerian Lingkungan Hidup dan Kehutanan dalam menentukan prioritas DAS mana yang akan direhabilitasi dan strategi rehabilitasi yang akan dilakukan.

Dari proses clustering dan uji validitas cluster yang telah dilakukan, didapatkan jumlah 4 cluster merupakan jumlah cluster optimal dengan nilai DBI rata-rata terendah yaitu 0.672778 atau lebih kecil 19,93% dari nilai DBI rata-rata untuk semua percobaan. artinya pengelompokan DAS dengan jumlah 4 (empat) cluster mempunyai tingkat kemiripan data dalam satu cluster yang relatif tinggi dibandingkan dengan jumlah lain.

Hasil clustering menunjukan bahwa DAS di cluster 3 dengan jumlah 332 DAS yang dominan tersebar Kab. Minahasa Selatan (24,7%) adalah DAS dengan tingkat kekritisian yang tinggi dibandingkan kelompok lain, Ini dikarenakan parameter karakteristik DAS yang ada di cluster 3 berada pada tingkat yang rendah jika dibandingkan dengan kelompok lain.

5.2 Saran

Pada penelitian berikutnya perlu dilakukan proses optimasi dalam penentuan centroid awal dalam clustering DAS dengan 4 cluster, ini dikarenakan pada algoritma K-modes sederhana penentuan centroid awal dilakukan secara random yang bisa menyebabkan perbedaan pada hasil clustering.

Penelitian selanjutnya juga perlu dilakukan untuk mengetahui segmen dan pola sebaran kelompok-kelompok DAS berdasarkan parameter karakteristiknya di provinsi lain di Indonesia, hal ini bertujuan agar karakteristik dan pola sebaran DAS di seluruh Indonesia bisa dipetakan.

[Halaman ini sengaja dikosongkan]

DAFTAR PUSTAKA

Achmad Yasid. “*Implementasi Automatic Clustering menggunakan Diffrencial Evolution dan CS Measure untuk Analisis Data Kemahasiswaan*”. *NERO. Indonesian Journal of Network Engineering and Research Operation*. 2014;2(1):47-52.

Akter, R, Chung Y. “*An Evulotionary Approach for Document Clustering*” . *IERI Procedia*. 2013;4(0): 370-375.

Chang, Kang-tsung. “*Introduction to Geographic Information System*”. McGraw-Hills New York. 2002.

Davies, D.L and Bouldin, D.W “*Cluster Separation Meassure*”. *IEEE Transaction on Pattern, Analysis and Machine Intelligence* 1(2):95-104. 1979.

Dephut, “*P.32/Menhut-II/2009 tentang Tata cara penyusunan rencana teknik rehabilitasi hutan dan lahan Daeah Aliran sungai*”. Jakarta, 2009.

Dephut, “*P.25/Menhut-II/2010 tentang Pedoman Penyelenggaraan Rehabilitasi Hutan dan Lahan*”. Jakarta, 2010.

Elma Hot, Vesna Popovic, *Soil Data clustering by using k-means and fuzzy k-means, Telecominication Forum (TELFOR)*, Belgrade.2015:890-893.

Handayani, Vidya. “*Analisis Clustering Menggunakan Algoritma K-Modes*”. Fakultas Teknik Informatika, Universitas Telkom . 2010.

Handoko, Slamet. “*Sistem Informasi Geografis Berbasis Web untuk Pemetaan Sebaran Alumni Menggunakan Metode K-Means*”. Magister Sistem Informasi Universitas Diponegoro, 2012.

Harianja, Hendri. “*Visualisasi K-Means Clustering pada Data Potensi Pertanian Desa di Bogor Menggunakan Mapserver*”. Departemen Ilmu Komputer. Institut Pertanian Bogor. 2008.

Irhamni, Firli, dkk. “*Optimalisasi Pengelompokan Kecamatan berdasarkan Indikator Pendidikan menggunakan Metode Clustering dan Davies-Bouldin Indexz*”, Seminar Nasional Sains dan Teknologi Universitas Muhamadiyah Jakarta, 12 November 2014.

Jingdong T, Rujing W, Bingyu S. “*Visualizing the Result of Fuzzy Clustering using GIS. Second International Workshop on Knowledge Discovery and Data Mining*”. Moscow. 2009: 267-271.

Liying Zhang, Hongge Hu. “*The Study of Synthesis Geographic Information on Fuzzy C-Means Clustering Algorithm*”. *International Conference on Consumer Electronic, Communication adn Network (CECnet)*. XianNing. 2011; 2804-2807.

Prasetyo, Eko, "Data Mining Mengolah Data menjadi Informasi dengan Matlab," Andi Yogyakarta, 2014.

Rencher ,Alvin C.. ” *Method of Multivariate Analysis*”. Brigham Young University, USA, 2002.

R.J. Kuo, Erma S, A Yasid. “*Automatic Clustering Combining Differential Evolution Algorithm and K-means Algorithm*”. *Proceeding of the Institute Engineers Asian Confrence*. Taipei.2013: 1207-1216.

S Khan, S Kant. “*Computation of Initial Modes for K-Modes Clustering Algorithm using Evidence Accumulation*”. *IJCAI*. Hyderabad-India. 2007; 2784-2789.

Santosa B, "Data Mining : Teknik Pemanfaatan Data untuk keperluan Bisnis," Graha Ilmu-Yogyakarta, 2007.

Sunyoto, Andi. “ *Integrasi PHP-MySQL dengan Google Maps APP*”. Jurnal DASI STMIK AMIKOM Vol.11 No.1 2010, Yogyakarta.

Supranto. “*Analisis Multivariat Arti dan Interpretasi Edisi Pertama*”. Rineka Cipta, Jakarta. 2004.

Surtiani, Yohan, B, Lilin. “*Evaluasi Rehabilitasi Hutan dan Lahan (RHL) di Daerah Aliran Sungai (DAS) Juwana pada Kawasan Gunung Muria Kabupaten Pati*”, Jurnal Pembangunan Wilayah Kota UNDIP Vol.11 No.1 2015:117-128,Semarang.

Susijanto T. Rasmana,Yoyon K Suprpto, I Kettut E, *Color Clustering in the Metal Inscription Images Using ANFIS Filter*.TELKOMNIKA. Indonesian Journal of Electrical Engineering. 2013;11(3): 529–536.

Tan , et.al. “ *Introduction to data mining*”. Pearson Addison Wesley Boston, 2006.

Tian, BAI, et.al “*A Global K-Modes Algorithm form Clustering Categorical Data*”, Chinese Journal of Electronics, Vol.21 July, 2012

Wu, X and Kumar, V.”*The Top Ten Algorithm in Data Mining*”. CRC Press Taylor & Francis Group. 2009.

Yuhefizar, Budi S,I Ketut E, Yoyon K Suprpto, *Combination of Cluster Method for Segmentation Web Visitors*.TELKOMNIKA. Indonesian Journal of Electrical Engineering. 2013;11(1): 207–214.

Yuhefizar, Budi S,I Ketut E, Yoyon K Suprpto, *Two Level Clustering Approach for Data Quality Improvement in Web Usage Mining*. *Journal of Theoretical and Applied Information Technology*. 2014:62(2):404-409.

BIOGRAFI PENULIS



Muhammad Farid Fahmi, lahir pada tanggal 20 Februari 1984 di Gresik, Jawa Timur. Setelah tamat dari SMA Semen Gresik pada tahun 2002, Penulis kemudian melanjutkan pendidikan S1 di Jurusan Teknik Informatika Fakultas Teknik Universitas Trunojoyo dan lulus tahun 2007. Penulis pernah bekerja di PT. Sumber Alfaria Trijaya, Tbk dan PT. Midi Utama Indonesia, Tbk sebagai IT Support sebelum di terima sebagai Pegawai Negeri Sipil (PNS) di Balai Penelitian Kehutanan Manado, Kementerian Kehutanan pada tahun 2010. Pada tahun 2014 penulis mendapatkan beasiswa Magister Telematika *Cief Information Officer* (CIO) dari Kementerian Komunikasi dan Informatika (Kemekominfo) di Jurusan Teknik Elektro Fakultas Teknologi Industri Institut Teknologi Sepuluh Nopember Surabaya. Saat ini, penelitian yang telah dilakukan oleh penulis lebih cenderung ke bidang Data Mining dan Sistem Cerdas. Penulis bisa dihubungi melalui email : muhammad.farid29@gmail.com .